

147838

Schreiber, David

From: Ramirez, Delia
Sent: Thursday, March 10, 2005 4:09 PM
To: Schreiber, David
Subject: case 10/034,985

Hi,

I would like to request the following search : SEQ ID NO:1 and 2 in the protein databases (commercial and interference).

Thank you,

Delia M. Ramirez, Ph.D.
Patent Examiner
Recombinant Enzymes-Art Unit 1652
USPTO
400 Dulany Street, Remsen Bldg., 2D74, Mail room 2C70
Alexandria, VA 22314
(571) 272-0938
delia.ramirez@uspto.gov

THIS PAGE BLANK (USPIC,

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 11, 2005, 23:36:45 ; Search time 183.5 Seconds
(without alignments)
4756.237 Million cell updates/sec

Title: US-10-034-985-1
Perfect score: 2349
Sequence: 1 atgaagcgatcttaatcc.....atcacatcacatcactaa 1323

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1396920 seqs, 329844858 residues

Total number of hits satisfying chosen parameters: 2793840

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p model -DEV=xlp
-O=/cgn2_1/USPTO_spool_P/US10034985/runat_10032005_164243_12058/app_query.fasta_1.1479
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US10034985 @CGN 1.1 206 @runat_10032005_164243_12058
-NCPU=6 -ICPU=1 -NO_MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRAD=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pcp:
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pcp:
3: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp:
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pcp:
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pcp:
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pcp:
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pcp:
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pcp:
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pcp:
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pcp:
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pcp:
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pcp:
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pcp:
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pcp:
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pcp:
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pcp:
17: /cgn2_6/ptodata/1/pubpaa/US10E_NEW_PUB.pcp:
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pcp:
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pcp:
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pcp:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

* Query
Result

No.	Score	Match	Length	DB	ID	Description
1	2302	98.0	440	9	US-09-777-566A-2	Sequence 2, Appli
2	2302	98.0	440	9	US-09-866-379-2	Sequence 2, Appli
3	2302	98.0	440	14	US-10-034-985-2	Sequence 2, Appli
4	2302	98.0	440	15	US-10-430-356-2	Sequence 2, Appli
5	2302	98.0	440	15	US-10-601-313-2	Sequence 2, Appli
6	2235	95.1	432	9	US-09-866-379-8	Sequence 8, Appli
7	2235	95.1	432	14	US-10-156-660-4	Sequence 4, Appli
8	2235	95.1	432	15	US-10-282-122A-43351	Sequence 43351, A
9	2235	95.1	432	15	US-10-601-319-8	Sequence 8, Appli
10	2235	95.1	432	16	US-10-472-317-41	Sequence 41, Appli
11	2221	94.6	432	15	US-10-284-962-3	Sequence 3, Appli
12	2221	94.6	432	15	US-10-284-962-14	Sequence 14, Appli
13	2221	94.6	432	14	US-10-266-041-1	Sequence 1, Appli
14	2197	93.5	432	15	US-10-284-962-5	Sequence 5, Appli
15	2188	93.1	436	14	US-10-156-660-2	Sequence 2, Appli
16	2171	92.4	432	15	US-10-601-319-10	Sequence 10, Appli
17	2165	92.2	432	14	US-10-021-723A-16	Sequence 16, Appli
18	2141	91.1	430	9	US-09-866-379-10	Sequence 10, Appli
19	2070	88.1	412	14	US-10-334-672-1	Sequence 1, Appli
20	2070	88.1	412	14	US-10-334-672-1	Sequence 1, Appli
21	2066	88.0	410	14	US-10-021-723A-13	Sequence 13, Appli
22	2066	88.0	410	14	US-10-021-723A-15	Sequence 15, Appli
23	960.5	40.9	441	14	US-10-021-723A-12	Sequence 12, Appli
24	943.5	40.2	441	15	US-10-282-122A-77792	Sequence 77792, A
25	935.5	39.8	441	14	US-10-021-723A-2	Sequence 2, Appli
26	932.5	39.7	420	14	US-10-021-723A-4	Sequence 4, Appli
27	760	32.4	476	14	US-10-021-723A-10	Sequence 10, Appli
28	611.5	26.0	318	14	US-10-021-723A-14	Sequence 14, Appli
29	476.5	20.3	409	14	US-10-021-723A-8	Sequence 8, Appli
30	472.5	20.1	421	14	US-10-021-723A-6	Sequence 6, Appli
31	195	8.3	426	15	US-10-257-174-44	Sequence 44, Appli
32	195	8.3	426	15	US-10-343-357-6	Sequence 6, Appli
33	126	5.4	3074	15	US-10-282-122A-65531	Sequence 65531, A
34	121.5	5.2	19723	15	US-10-084-846A-5	Sequence 5, Appli
35	121	5.2	465	16	US-10-662-914-10	Sequence 10, Appli
36	120.5	5.2	19695	15	US-10-084-846A-3	Sequence 3, Appli
37	119	5.1	439	14	US-10-062-846-3	Sequence 3, Appli
38	119	5.1	439	16	US-10-776-104-3	Sequence 106, App
39	119	5.1	440	15	US-10-442-538-106	Sequence 128, App
40	119	5.1	440	15	US-10-442-538-128	Sequence 151, App
41	119	5.1	449	15	US-10-442-538-151	Sequence 12, Appli
42	119	5.1	449	16	US-10-776-104-12	Sequence 8, Appli
43	119	5.1	465	13	US-10-083-452-8	Sequence 78, Appli
44	119	5.1	465	14	US-10-062-848-78	
45	119	5.1	465	14	US-10-062-848-78	

ALIGNMENTS

RESULT 1
US-09-777-566A-2
; Sequence 2, Application US/09777566A
; Patent No. US20010055788A1
; GENERAL INFORMATION:
; APPLICANT: DIVERGA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-6
; CURRENT APPLICATION NUMBER: US/09/777,566A
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: US 05/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.0

SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n is any nucleotide
US-09-777-566A-2

Alignment Scores:

Pred. No.: 5,44e-205 Length: 440
Score: 2302.00 Matches: 440
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.00% Indels: 0
DB: 9 Gaps: 0

US-10-034-985-1 (1-1323) x US-09-777-566A-2 (1-440)

QY 1 ATGAAAGCGATCTAATCCCATTTTATCTCTCTGATTCGGTTAAACCCCGCAATCTGCA 60
DB 1 MetLysAlaIleuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTGCTCAGAGTGAAGCGGAGCTGAAGCTGGAAGTGTGGTGTGATGTGTCAGTCGTCATGGT 120
DB 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerA-gHisGly 40
QY 121 GTGCTGCTCAACAAAGGCCAGCAACTGATGAGATGTACCCCGACGATGGCCA 180
DB 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGGCGGTAAACATGGGTGGCTGACACCGCGGGTGGTGTGAGCTAATCGCTATCTC 240
DB 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyLeu 80
QY 241 GGACATTACCAACCCAGCGCTGCTGATGCGACAGGATGTGCGCAAAAGGGGTGCGCG 300
DB 81 GlyHisTyrglnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGGTCAGTTCGGATATTGCTGATGTGACGAGCGTACCCGTTAAACAGGGCAA 360
DB 101 GlnSerGlyGlnValAlaIleAlaSerValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCCCGCGCGTGGCACCTGACTGTGCAATAACCGTACATACCCAGGACGATACG 420
DB 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGATCCGTTATTATTCCTCTAAACTGGCGTTTCCCACTGGATAACGCG 480
DB 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AACGTGACTGACGATCTCAGACAGGCGAGGAGGTCATTTGCTGACTTTACCGGCAT 540
DB 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAACCGGTTTCGGACCTGACCGGTGCTTAATTTCCGCAATCAACTCTGTC 600
DB 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTGAGAACAGGACGAAGCTGTTTCATTAAACGAGGATTAACCATCGGAATC 660
DB 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGGTGACGCGCAATGTCTAATTAACCGGTGCGGTAAAGCTCGCATCAATCTGACG 720
DB 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGTATTTCTCTGCAACAGCACAGGATCCGCGAGCGGGTGGGAAGCATCAC 780
DB 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTACACACGAGTGGAAACACCTTGCTAAGTTTGATTAACGCGCAATTTTCTCTACAA 840

DB 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyLeuLeuGln 280
QY 841 CGCACCCAGAGGTTCGCCGCGAGCCGCCCGTATTGATTTGATTCATGGCAGCG 900
DB 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
QY 901 TTGACGCCCATCCACGCAAAACAGGCGGTATGGTGTGACATTAACCATTCAGTACTG 960
DB 301 LeuThrProHisProGlnLysGlnAlaTyrglyValThrLeuProThrSerValLeu 320
QY 961 TTTATTGCGGACACGATACCTAATCTGGCAAAATCTCGGCGCGCACTGGAGCTCACTGG 1020
DB 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTCCTCCGTCAGCGGATACACGCCGCGAGGTGGTGAACCTGGTGTGAAACGCTGG 1080
DB 341 ThrLeuProGlyGlnProAspAsnThrProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTCGGCTAAGCCATAACAGCCAGTGTGATTCAGGTTTCGCTGCTTCCAGACTTTACAG 1140
DB 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATCGGTGATAAAACGCGCTGTCTAATAATACGCGCCCGAGAGGTGAAACTGACC 1200
DB 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlyGluValLysLeuThr 400
QY 1201 CTGCGAGGATGTCAAGAGCGAAATCGCAGGCGCATGTTCGTTGGCAGGTTTTACGCAA 1260
DB 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATGAAGCAGCATACCGCGGTGAGTTTGAGATCTCATCACCATCACCATCAC 1320
DB 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440

RESULT 2

US-09-866-379-2
; Sequence 2, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-866-379-2

Alignment Scores: 5,44e-205 Length: 440
Pred. No.: 2302.00 Matches: 440
Score: 100.00% Conservatives: 0
Percent Similarity: 100.00%


```
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.00% Indels: 0
DB: 9 Gaps: 0

US-10-034-985-1 (1-1323) x US-09-866-379-2 (1-440)

Qy 1 ATGAAGCGATCTTAATCCCAATTTTATCTCTTCTGATTCGGTTAAACCCGCAATCTGCA 60
Db 1 MetlysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20

Qy 61 TTCGCTCAGAGTAGCGGAGCTCAAGCTGGAAGTGTGGTATTGTCAGTCGTCATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40

Qy 121 GTGCGTCTCCAAACCAAGCCACCAACTGATGAGGATGTACCCGAGAGCATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60

Qy 181 ACCTGGCGGTAAACCTGGTGGCTGACACCGCGGTGTGAGCTTAATCGCCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80

Qy 241 GGACATTACCAACGCCAGCTCTGCTAGCCAGCGATTGTCGCGAAAGGGCTGCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100

Qy 301 CAGTCTGCTCAGTCCGCGATTATTGCTGATGTCACGAGCGTACCCGTAAACACAGCGAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120

Qy 361 GCCTTCGCGCGCGGTGGCACTGCTGCAATAACCGTACATACCCAGCGCAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140

Qy 421 TCCAGTCCCGATCCGTTATTATCTTAATCTTAATAAATCGCGTTCGCACTGGATAACGG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160

Qy 481 AACGTGACTAGCGGATCCTCAGCAGGCGAGGGTCAATTGTGACTTTTACCGGCGAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180

Qy 541 CGGCAACCGCGGTTTCGCAACTCGAATCGGCTGCTTAATTTTCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200

Qy 601 CTTAAACGTGAGAACAGCAGCAAGCTGTTCATTACGAGCGATTAACCTCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220

Qy 661 AAGGTGAGCGCGCAATGCTCTCATTAACCGGTGCGGTAAAGCCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240

Qy 721 GAGATATTCTCTCGCAACAGCAGCGAATGCGGAGCGGGGTGGGAAGGATCAC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260

Qy 781 GATTACACAGTGGAGACCTTCTGCTAAGTTGATTAACCGCGCAATTTTATTTCTCTCAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280

Qy 841 CGCAGCGCAGAGTTGCCGCGCGCCACCGCTTATTTGATTTGATTCATCGCAGCG 900
Db 281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300

Qy 901 TTGACGCGCCATCCACCGCAAAACAGCGGTATGCTGATTAACCCACTTTCAGTACTG 960
Db 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320

Qy 961 TTTATTCGCGACACGATCTAAATCTGCGCAAAATCTGCGCGCGCAGCTGAGACTCAATGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
```

```
Qy 1021 ACCTTCCCGTCCAGCCGATAAACACGCCCCAGGTGCTGAACCTGGTGTGAACTGGT 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360

Qy 1081 CGTCCGCTAAGCATAACACGCGAGTCGATTACGTTTCGGTTCGCTTCCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380

Qy 1141 CAGATCGCTGATAAAACGCCGCTGTCTAATAATACGCCGCCGAGAGGTGAAACTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400

Qy 1201 CTGCGAGGATGTGAAGAGCGAAATGCGCAGGGCATGTTCGTTGGCAGGTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420

Qy 1261 ATCGTGATCAAGCAGCCATACCGCGGTGCGAGTTGAGATCTCATCACCATCACCATC 1320
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis 440
```

RESULT 3

```
US-10-034-985-2
; Sequence 2, Application US/10034985
; Publication No. US20030049815A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/034,985
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-034-985-2
```

```
Alignment Scores:
Pred. No.: 5,44e-205 Length: 440
Score: 2302.00 Matches: 440
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.00% Indels: 0
DB: 14 Gaps: 0
```

US-10-034-985-1 (1-1323) x US-10-034-985-2 (1-440)

```
Qy 1 ATGAAGCGATCTTAATCCCAATTTTATCTCTTCTGATTCGGTTAAACCCGCAATCTGCA 60
Db 1 MetlysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20

Qy 61 TTCGCTCAGAGTAGCGGAGCTCAAGCTGGAAGTGTGGTATTGTCAGTCGTCATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40

Qy 121 GTGCGTCTCCAAACCAAGCCACGCAACTGATGAGGATGTACCCGAGAGCATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60

Qy 181 ACCTGGCGGTAAACCTGGTGGCTGACACCGCGGTGTGAGCTTAATCGCCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
```

```
Qy 241 GGACATTACCAACGCCAGCGCTGTGTAGCCGACGATTGTGGCGAAAAAGGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Qy 301 CAGTCTGGTCAGTCGGGATTATTGCTGATGTACAGAGGTACCCGTAAACAGCGGAA 360
Db 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Qy 361 GCCTTCGCCCGCGGTGGACCTGACTGTGCATTAACCGTACATACCCAGGCGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Qy 421 TCCAGTCCCGATCGTTATTATTCCTCTAAACTCGCGTTGCGCAACTGGATAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Qy 481 AACTGTACTGACGGATCTCAGAGGCGAGGAGGTCAATTGCTGACTTTTACCGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Qy 541 CGGCAAAACGGGTTTCGGCACTGGAACGGGTGCTTAATTTCCGCAATCAAACTTCTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Qy 601 CTTAAACGTGAGAAACAGGACGAAGCTGTTTCAATTAACGAGGCAATTACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Qy 720 AAGGTAGCGCGCAANTGCTCTTAATTAACCGGTGCGTAAAGCTTCGGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Qy 780 GAGATATTCTTCGCAACAGCACGAGATCGCGAGCGCGGTGGGAAGGATCAC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrPglyArgIleThr 260
Qy 840 GATTACACACGAGTGGAAACACCTTCTTAAGTTTGATTAACGCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Qy 900 CGCACCGCAGAGTTGCGCAGCGCGCGCACCCCGTTTATTTGATTTGATCATGCGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Qy 960 TTGACGCGCCATCACCGCAAAACAGCGCATGTGTGATTCATACCACCTTCAGTACTG 960
Db 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Qy 1020 TTTATTCCCGGACACGATACCTAATCTGGCAATCTCGCGCGCGCATCTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
Qy 1080 ACGTTCCCGTACGCGGATACACCGCGCAGGTGTGAACTGGTTTGAACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Qy 1140 CGTCGGTAAACGATACAGCCAGTGTGATTCAGGTTTCGTTGCTTCCAGCTTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Qy 1200 CAGATGGGTGATAAAACGCCCTCTCATTAATACGCCGCCCGGAGAGGTGAACTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Qy 1260 CTGCGCAGATGTGAAGACGGAATTCGCGAGGCGATGTGTTGTTGGCAGGTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Qy 1320 ATCGTGAATGAACACCATACCGCGTGCAGTTTGTAGATCTCATCCATCCATCCATCAC 1320
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
```

```
US-10-430-356-2
; Sequence 2, Application US/10430356
; Publication No. US20030232041A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/430,356
; CURRENT FILING DATE: 2003-05-05
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-430-356-2
```

```
Alignment Scores:
Pred. No.: 5,44e-205 Length: 440
Score: 2302.00 Matches: 440
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.00% Indels: 0
DB: 15 Gaps: 0
```

US-10-034-985-1 (1-1323) x US-10-430-356-2 (1-440)

```
Qy 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTCTGATTCCTGTTAAACCCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Qy 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGTGATGTGTCAGTCTGTCATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Qy 121 GTGCGTCTCAACCAAGGCCACGCACTGATCAGAGTGTACCCAGACGATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Qy 181 ACCTGGCCCGTAAAACTGGGTTGGCTGACACCCGCGNGGTGGTGAAGCTAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Qy 241 GGCATTTACCAACGCCAGCGCTCTGGTAGCCGAGGATTCGTCGGGAAAAAGGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Qy 301 CAGTCTGGTCAGTCGGGATTATTGCTGATGTGCAGAGGTACCCGTAACAGAGCGGAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Qy 361 GCCTTCGCCCGCGGTGGCACCTGACTGTGCAATTAACCGGTACATACCCAGGCGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Qy 421 TCCAGTCCCGATCGTTATTATTCCTCTAAACTCGCGTTGCGCAACTGGATAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Qy 481 AACTGTACTGACGGATCTCAGAGGCGAGGAGGTCAATTGCTGACTTTTACCGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
```

```

Qy 541 CGGCACCGCGTTTCGCGAACTGGAACGGGTGCTTAATTTTCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Qy 601 CTTAAACGTGAGAACAGGACGAAAGCTGTTTCATTAACGAGGAGCATTAACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Qy 661 AAGTGAGCGCCGACATGCTCTAATTAACCGGTGCGGTAAAGCTTCGATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Qy 721 GAGATATTCTCTCGACACAGCAGGATCGCGGCGCGGTGGGAGGATCACC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Qy 781 GATTACACACAGTGGAAACACCTCTTAAGTTTGATTAACGCGCAATTTTCTCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Qy 841 CGCACGCGAGAGTTGCCCGCAGCGCCGACCCCGTTATTGGATTGATCANGCGACG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Qy 901 TTGACGCGCCCATCCACGCAAAACAGCGGTATGCTGTGACATTACCACCTTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Qy 961 TTTATTCCGCGACACGATACTAACTCGCAATCTCGCGCGGACCTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
Qy 1021 ACGCTTCCCGGTACGCGGATTAACACGCGCGAGTGTGTAACCTGTTTGAACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Qy 1081 CGTCGGCTAACGATACACGACGATGATTCAGTTTCCGTTCCGAGACTTTACG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Qy 1141 CAGATGCTGATAAACGCGCTCTCATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Qy 1201 CTGCGCAGGATGTAAGACGCAAAATCGCGAGGCGCATGTGTTGTTGCGCAGGTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Qy 1261 ATCTGTAATGAAGCACGATACCGCGGTGAGTTGAGATCTCATCACTACCATCAC 1320
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440

```

RESULT 5

US-10-601-319-2

; Sequence 2, Application US/10601319

; Publication No. US20040091968A1

; GENERAL INFORMATION:

; APPLICANT: Short, Jay M.

; APPLICANT: Kretz, Keith A.

; APPLICANT: Gray, Kevin A.

; APPLICANT: Barton, Nelson Robert

; APPLICANT: Garrett, James B.

; APPLICANT: O' Donoghue, Eileen

; APPLICANT: Mathur, Eric J.

; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING

; FILE OF INVENTION: AND USING THEM

; FILE REFERENCE: 09010-029011

; CURRENT APPLICATION NUMBER: US/10/601.319

; CURRENT FILING DATE: 2003-06-20

; PRIOR APPLICATION NUMBER: US 09/866,379

; PRIOR FILING DATE: 2001-05-24

; PRIOR APPLICATION NUMBER: US 09/580,515

; PRIOR FILING DATE: 2000-05-25

```

; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-601-319-2

```

Alignment Scores:

```

Pred. No.: 5,44e-205 Length: 440
Score: 2302.00 Matches: 440
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.00% Indels: 0
DB: 15 Gaps: 0

```

US-10-034-985-1 (1-1323) x US-10-601-319-2 (1-440)

```

Qy 1 ATGAAGCGATCTTAATCCCATTTTATCTCTTCTGATTCGTTAACCCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Qy 61 TTGCTCAGAGTGAAGCGGAGCTGAAGCTGGAAGTGTGTTGATTGTCAGTCTCATGTT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Qy 121 GTGCGTCTCAACACGAGCGACGCACTGATGCAGGATGTCACCCAGACGCAATGSCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Qy 181 ACCTGGCGGTAAACATGGGTTGGCTGACACCGCGNGGTGGTGAAGCTAAATCGCCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Qy 241 GACATTAACCAACGCGAGCTCTGTGTAGCCGACGAGATTGTCGGCGAAAAAGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Qy 301 CAGCTCTGTCAGCTCCGATTATTGCTGATGTCGAGCGGTACCCCTTAAACAGGCGAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Qy 361 GCCTTCGCGCGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Qy 421 TCCAGTCCCGATCCGTTATTTAATCTCTAAAAAATCGCGCTTTCGCAATCAAACTTGTGC 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Qy 481 AACGTGACTACGCGATCTCTCAGCGGCGAGGAGGTCAATTCGTGACTTTTACCGGCGAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Qy 541 CGCAAAACGGCGTTTCGCAACTGGAACGGGTGCTTAATTTTCCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Qy 601 CTTAAACGTGAGAACAGGACGAAAGCTGTTTCATTAACGAGGAGCATTAACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Qy 661 AAGTGAGCGCCGACATGCTCTAATTAACCGGTGCGGTAAAGCTTCGATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240

```

QY 721 GAGATATTTCTCTGCAACAGCACAGGAATCCGCGAGCCGGGTGGGAAGGATCACC 780
Db 241 GluilePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgileThr 260
QY 781 GATTACACAGTGGAAACACCTTGCTAAGTTTGATACAGCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
QY 841 CGACGCGCAGAGTGGCCGCGAGCCGCGCCACCCCGCTTATGATTTGATCATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuileMetAlaAla 300
QY 901 TTGACGCGCCATCACCGCAAAACAGCGGTATGCTGATACATTACCACCTTCAGTACTG 960
Db 301 LeuThrProHisProProGlnLysGlnAlaArgLysValThrLeuProThrSerValLeu 320
QY 961 TTTATTCCCGACACGATACCTAATCTGCAAAATCTCGCGCGCGCAGCTGAGCTCAACTGG 1020
Db 321 PheileAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTCCCGGTACGCGGATACACCGCCGAGGTGGTGAACCTGGTTTGAACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTCGGCTAAGCGATACACGAGTGGATTCAGGTTTCGCTGCTTCCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATGCGTGATAAAACCGCGCTCTCATTAATACCGCGCGCGAGAGGTGAAACTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGCGCAGATGTGAAGCGCAATGCGCAGCGCATGTTGCTGGCAGGTTTACCCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGATGAAGCACCATACCGCGCTGAGTTTGAGATCTCATACCATCCATCAC 1320
Db 421 IleValAsnGluAlaArgileProAlaCysSerLeuArgSerHisHisHisHisHis 440

RESULT 6

US-09-866-379-8
; Sequence 8, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 432
; TYPE: PRF
; ORGANISM: Escherichia coli

US-09-866-379-8

Alignment Scores:
Pred. No.: 9,26e-199 Length: 432
Score: 2235.00 Matches: 430
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 2
Query Match: 95.15% Indels: 0
DB: 9 Gaps: 0

US-10-034-985-1 (1-1323) x US-09-866-379-8 (1-432)

QY 1 ATCAAAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGTTAAACCCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuileProPheLeuSerLeuLeuileProLeuThrProGlnSerAla 20
QY 61 TTGCTCAGAGTGAAGCGGAGCTGAAGCTGGAAGTGTGGTATTGTCTAGTCCTCATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
QY 121 GTGCGTGTCTCCACCAAGCCACGCAACTGATGCAGATGTCAACCCAGACGATGSCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGCGCGTAAACCTGGGTTGGCTGACACCCCGNGGTGGTGGAGCTAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuileAlaTyrLeu 80
QY 241 GGCATTTACCAACCGCAGCGTCTGGTAGCGAGGATTGCTGCGGAAAAAAGGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGCTGTGTGAGTCCGCTATTATTGCTGATGTCGACGAGCGTACCGTAAACAGGCGAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCCGCGCGCGGTGGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGATCCGTTATTATCTCTAATAAAGTGGCGTTGCGCAATCTGACTTTTACCGG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AAGTGTGACTGACCGGATCTCTCAGCAGCGAGGAGGTCAATTGCTGACTTTTACCGG 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAAAACCGCGTTTCGCGAACTGGAAACCGGCTCTTAATTTTCCGCAATCAAACTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAAACCGGTCGTAAGCTTACCATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTGTGAGCGCGCACATGTCTCATTAAACCGGTCGTAAGCTTACCATC 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTTCTCTGCAACAGCACAGGGAATCCGCGAGCCGGGTGGGAAGATCACC 780
Db 241 GluilePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgileThr 260
QY 781 GATTACACAGTGGAAACACCTTGCTAAGTTTGATACAGCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
QY 841 CGCAGCGCGCAGGTTGCGCGCAGCGCCACCCCGTTATTGATTTGATCATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuileLysThrAla 300
QY 901 TTGACGCGCCATCCACCGCAAAACAGGCGGTATGGTGTGACATTACCCACTTCAGTACTG 960

301	LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu	320
961	TTTATTCCCGGACACGATACTAATCTCGCAAACTCTCGCGCGCACGTGGAGCTCAACTGG	1020
321	PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp	340
1021	ACGCTTCCCGGTACGCCGGATAACAACGCCCGCAGGTGGTGAACCTGGTTTGAACGCTGG	1080
341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp	360
1081	CGTCGGCTAAGCGATAACAGCCAGCATGGATTACAGTTTCGTGGTCTTCCAGACTTTACAG	1140
361	ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln	380
1141	CAGATCGTGATAAACCCCGCTGTCTAATAATACGCCGCCGAGAGGTGAAACTGACC	1200
381	GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr	400
1201	CTGCGCAGGATGTGAAGACGGAATATGCCAGGCGCATGTTCGTGGCAGGTTTACGCCAA	1260
401	LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln	420
1261	ATCGTGAATGAAGCACCGCATACCGCGCTGCAGTTTG	1296
421	IleValAsnGluAlaAraGlieProAlaCysSerLeu	432

RESULT 7

```

RES001
US-10-156-660-4
; Sequence 4, Application US/10156660
; Publication No. US20030103958A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson R.
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM
; TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-029007
; CURRENT APPLICATION NUMBER: US/10/156,660
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 432
; TYPE: PRF
; ORGANISM: Escherichia coli
US-10-156-660-4

```

Alignment Scores:				
Pred. No.:	9,266-199	Length:	432	
Score:	2335.00	Matches:	430	
Percent Similarity:	99.54%	Conservative:	0	
Best Local Similarity:	99.54%	Mismatches:	2	
Query Match:	95.15%	Indels:	0	
DB:	14	Gaps:	0	
US-10-034-985-1 (1-1323) x US-10-156-660-4 (1-432)				
Qy	961	TTTATTGCGGACAGCATACTAAATCTGCAAAATCTCGCGCGGCACTGGAGCTCAACTGG	1020	
Db	321	PhEiLeAlaGlyHiAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp	340	
Qy	1021	ACGCTTCCCGGTACGCGCGATAAACACGCCCGACGTTGGTGAACCTGGTGTGTTTGAACCGCTGG	1080	
Db	341	ThrLeuProGlyGlnProAspAsnThrProGlyGlyGluLeuValPheGluArgTrp	360	

QY 1081 CGTCGGCTAAGCGATACACCGCAGTGGATTACAGGTTTCGCTGCTTCCAGACTTACAG 1140
 Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 QY 1141 CAGATGGTGAATAAAGCCGCTGCTCAATTAATACCGCCCGGAGAGGTGAACCTGACC 1200
 Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 QY 1201 CTGGCAGGATGTGAAGACGCAATTCGCGAGGATGTGTTGCTGGCAGGTTTACGCAA 1260
 Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 ATCGTGAATGAAGACGATACCGGCTGCGATTG 1296
 Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

RESULT 8

US-10-282-122A-43351
 ; Sequence 43351, Application US/10282122A
 ; Publication No. US20040029129A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Liangsu
 ; APPLICANT: Zamudio, Carlos
 ; APPLICANT: Malone, Cheryl
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Kari
 ; APPLICANT: Zyskind, Judith
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trawick, John
 ; APPLICANT: Carr, Grant
 ; APPLICANT: Yamamoto, Robert
 ; APPLICANT: Forsyth, R.
 ; APPLICANT: Xu, H.

; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A

; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/230,335

; PRIOR FILING DATE: 2000-09-06

; PRIOR APPLICATION NUMBER: 60/230,347

; PRIOR FILING DATE: 2000-09-09

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/267,636

; PRIOR FILING DATE: 2001-02-09

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 78614

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 43351

; LENGTH: 432

; TYPE: PRT

; ORGANISM: Escherichia coli

US-10-282-122A-43351

Alignment Scores:

Pred. No.: 9,26e-199 Length: 432
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 95.15% Indels: 0
 DB: 15 Gaps: 0

US-10-034-985-1 (1-1323) x US-10-282-122A-43351 (1-432)
 QY 1 ATGAAAGCATCTTAATCCCATTTTATCTCTTCTGATTCGTTAAACCCCGCAATCTGCA 60
 Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 QY 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGGTGAATGTCTAGTCTGCTATGGT 120
 Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
 QY 121 GTGCTGTCTCCAAAGCCAGCCAGCAACTGATGATCAGAGTGTCAACCCAGAGCGATGGCCA 180
 Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 QY 181 ACCTGCCCGGTAAACACTGGGTTGGCTTGACACCCGCGNGTGTGAGTAACTCGGCTATCTC 240
 Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
 QY 241 GGACATTACCAACGCCAGCGTCTGTGTAGCCGACGAGTTGTCGCGAAAGAGGCTGCCCG 300
 Db 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 QY 301 CAGTCTGTGTCAGTCCCGATTATTGCTGATGTCGACGAGGTACCCGTAAACAGCGGAA 360
 Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 QY 361 GCCTTCGCGCGCGGTGGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG 420
 Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 QY 421 TCCAGTCCCGATCCGTTATTTAATCTCTAAAACTGGCGTGGCTTTGCCAATCGATAACGCG 480
 Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 QY 481 AACGTGACTGACCGGATCTCTCAGCAGGCGAGGGGTCAATTGCTGACTTTACCGGGCAT 540
 Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 QY 541 CGGCAAAACGGCGTTTCGCAACTGGAAACGGGTCTTAATTTTCCGCAATCAAACTTGTCG 600
 Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 QY 601 CTTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATACGAGGAGTATCAATCGGAATC 660
 Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 QY 661 ACGTGAGCGCGACATGTCTCTAATACCGGTGCGGTAGCGCTCGCATCAATGCTGACG 720
 Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 QY 721 GAGATATTTCTCTGCAACAGCACAGGGAATCCGAGCGCGGGTGGGGAAGGATCACCC 780
 Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 QY 781 GATTTCACACCGAGTGAACACCTTTCCTAAGTTTCATAACCGCAATTTTATTTGCTACAA 840
 Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
 QY 841 CGCAGCGCAGAGTTGCGCGCAGCCCGCCACCCGTTATTGATTTGATTCATGCGCAGCG 900
 Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
 QY 901 TTGACGCCCCATCCACCGCAAAACAGGCGGTATGGTGTGATCATTTACCACTTCAGTACTG 960
 Db 301 LeuThrProHisProGlnLysGlnAlaTyrglyValThrLeuProThrSerValLeu 320
 QY 961 TTTATTGCGGACACGATACATACTATGCGCAAAATCTCGCGCGCGCATGAGAGCTCACTGG 1020
 Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 QY 1021 ACGTTCGCGGTGACGCCGATACACCGCCAGGTGGTGAACCTGGTGTGAACTGG 1080

Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTyr 360
Qy 1081 CGTCGGCTAAGCAGTAAACAGCAGTGTGATTCAGGTTTCGGTCTTCAGACTTACAG 1140
Db 361 ArgGluSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Qy 1141 CAGATGGTGATAAAGCCGCTGTCAATTAATACGCCGCCGAGAGGTGAACACTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Qy 1201 CTGCGCAGGATGTGAAGAGCGAAATGCGCAGGCGATGTTCGTTGGCAGGTTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Qy 1261 ATGCTGAATGAAGCAGCAGTACCGGCGTGCAGTTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

RESULT 9

US-10-601-319-8
; Sequence 8, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; TITLE OF INVENTION: AND USING THEM
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-601-319-8

Alignment Scores:

Pred. No.:	9,26e-199	Length:	432
Score:	2235.00	Matches:	430
Percent Similarity:	99.54%	Conservative:	0
Best Local Similarity:	99.54%	Mismatches:	2
Query Match:	95.15%	Indels:	0
DB:	15	Gaps:	0

US-10-034-985-1 (1-1323) x US-10-601-319-8 (1-432)

Qy 1 ATGAACGATCTTAATCCATTTTATCTTCTGATTCGGTTAAACCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Qy 61 TTGCTCAGATGAGCGGAGCTGAAGCTGGAAGTGTGGTGTGATGTGTCGTGATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40

Qy 121 GTGCGTGCTCCAAACCAAGGCCACGCAACTGATCAGGATGTCACCCAGACGATGSCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Qy 181 ACCTGGCCGCTAAACACTGGGTTGGCTGCACACCGCGNGGTGGTGAAGCTAAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
Qy 241 GGACATTACCAACGCCAGCGCTCTGGTAGCCGAGCGATTGCTGGCGGAAAAAGGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Qy 301 CAGCTGGTCCAGTCCCGATTATTGCTGATGTCGACGAGCGGTACCCGTAAACAGGCGAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Qy 361 GCCTTCGCCCGCGGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Qy 421 TCAGTCCCGATCCGTTATTTAATCTCTAAAACTGGCGTTTGGCCAACTGGATAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Qy 481 AACGTGACTGACGCGATCCCTCAGCAGCGCAGGAGGTCAATTGCTGACTTTTACCGGSCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyLysIleAlaAspPheThrGlyHis 180
Qy 541 CGGCAAAACGGCGTTTCGCAACTGGAACTGGGTCCTTAATTTTCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Qy 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAAACGCGGTAGCCCTCGCATCAATGCTACG 720
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Qy 720 AACGTGAGCGCCGACATGCTCATTAACCGGTGCGGTAGCCCTCGCATCAATGCTACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Qy 721 GAGATATTTCTCTGCAACAGCACAGGAAATCCGAGCGCGGTGGGAGGATCACC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Qy 781 GATTTCACACGAGTGAACACCTTTGCTAAGTTTGCATAACCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
Qy 841 CGCACCGCAGAGTTGCCCGCAGCGCCGCCCGCTTATGATTTGATTCATGCGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
Qy 901 TTCACGCCCATCCACGCAAAACAGCGGTATGTTGATACATTACCCACTTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
Qy 961 TTTATTGCCGACACGATACTAATCTGGCAAAATCTCGCGCGCGCAGCTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Qy 1021 ACGCTTCGCCGTCAGCCGGATAAACAGCCGCCAGGTGGTGAAGCTGGTGGTGGTAAACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Qy 1081 CGTCCGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTGCTTCCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Qy 1141 CAGATGCTGATAAAGCCGCGCTGTCTATTAAATACGCCCGCCGAGAGGTGAACACTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Qy 1201 CTGCGCAGGATGTGAAGAGCGAAATGCGCAGGCGATGTGTTGCTGTGGCAGGTTTTTACGCAA 1260

```
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATCAACGACGATACCGCGGTGCAGTTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
RESULT 10
US-10-472-317-41
; Sequence 41, Application US/10472317
; Publication No. US20040185562A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Incorporated
; TITLE OF INVENTION: Myo-Inositol Oxygenases
; FILE REFERENCE: 10829/003US1
; CURRENT APPLICATION NUMBER: US/10/472,317
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/US02/08404
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/277,148
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-472-317-41
Alignment Scores:
Pred. No.: 9,26e-199 Length: 432
Score: 2235.00 Matches: 430
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 2
Query Match: 95.15% Indels: 0
DB: 16 Gaps: 0
US-10-034-985-1 (1-1323) x US-10-472-317-41 (1-432)
QY 1 ATCAAGCGATCTTAATCCCATTTTATCTCTGATTCGGTTAAACCCCGCAATCGCA 60
Db 1 MetLysAlaIleuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTACGCGGAGCTGAAGCTGAAAGTGTGTGATTTGTCACTCGTCATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
QY 121 GTGGTGCTCCACCAAGGCCACGCAACTGATGACGAGATGTACCCACGACGATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGGCCGTTAAACTGGGTGGCTGACACCGCGNGGTGTGAGCTTAATCGCCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaIleLeu 80
QY 241 GGACATTACCAACGCCAGCGTCTGGTAGCCGACGAGATTGTCGCAAAAAGGGGTGCCCG 300
Db 81 GlyHisIleArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGGTCAGGTCGGGATATTGCTGATGTGACGAGGTACCCGTTAAACACGCGGAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCCCGCGGTGGACCTGACTGTGCAATACCGTACATACCCAGGCGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGCATCGTTATTATTAATCTCTAAAACTCGCGTTTGCCTGCAATACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AACGTGACTGACGGATCTTCAGACGGCGAGGGGTCAATTCGTGACTTTACCGGCAT 540
```

```
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGCAAAACGCGGTTTCGCAACATGGAACGCGTCTTAATTTTCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGAGCATATTCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AGGTGAGCGCCGACATGCTCTCAATTAACCGGTGCGGTAAAGCTCCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTCTCTCTCAACAGACAGGGAATCCCGAGCCGCGGTGGGGAAGGATCACCC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTACACAGTGGAAACCTTGTCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
QY 841 CGCACCGCAGAGGTTGCCCGCACCGCGCCACCCCGTATTGATTTGATTCATCGGACGC 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTCACGCCCCATCCACCGCAAAACAGGCGGTATGTTGATGATTCATCCACTTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyrglyValThrLeuProThrSerValLeu 320
QY 961 TTTATTGCGCGACACGATACTAATCTGGCAAAATCTCGGCGCGCAGCTGAGAGTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGCTTCCCGCTCAGCGGATACACGCGCCAGGTGGTGAACCTGTTTGAACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProGlyGlyGlyLeuValPheGluArgTrp 360
QY 1081 CQTGCGGTAAAGCCATAACACCGCATGAGTTCAGTTTCGCTGCTTCCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATCGGTGATAAAACGCGCTGTCTAATTAATACGCGCCCGGAGAGGTGAACCTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlyGluValLysLeuThr 400
QY 1201 CTGCGCAGGATGTCAAGACGGAATGCGAGGCGATGTTGCTTGGCAGGTTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATGAACGACGATACCGCGGTGCAGTTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
RESULT 11
US-10-284-962-3
; Sequence 3, Application US/10284962
; Publication No. US20030206913A1
; GENERAL INFORMATION:
; APPLICANT: Weibel, Douglas M.
; APPLICANT: Orr, Donald E.
; APPLICANT: Ruch, Frank E.
; TITLE OF INVENTION: PHYTASE-CONTAINING ANIMAL FOOD AND METHOD
; FILE REFERENCE: 834460-71725
; CURRENT APPLICATION NUMBER: US/10/284,962
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/335,303
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 432
; TYPE: PRT
```



```
; ORGANISM: Escherichia coli
US-10-284-962-3

Alignment Scores:
Pred. No.:      1.86e-197      Length:      432
Score:          2221.00      Matches:      428
Percent Similarity: 99.07%      Conservative: 0
Best Local Similarity: 99.07%      Mismatches: 4
Query Match:      94.55%      Indels:      0
DB:              15          Gaps:      0

US-10-034-985-1 (1-1323) x US-10-284-962-3 (1-432)

Qy 1 ATGAACCGATCTTAATCCCATTTTATCTCTCTGATTCGGTTAAACCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Qy 61 TTCGCTCAGATGAGCGGAGCTGAAGCTGGAAGTGTGTGATTCAGTCGTCATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Qy 121 GTGCGTCTCAACCAAGGCCACCAACTGATGCAGGATGTCACCCAGAGCATGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Qy 181 ACCTGGCCGTTAAATCTGGTGTGCTGACACCGCGNGGTGTGAGCTAAATCGCCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyLeu 80
Qy 241 GGACATTACCAACCCGAGCGTCTGGTAGCCGACGGATGCTGGCGAAAGGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Qy 301 CAGTCTGTCAGTCGCGATATTCTGATGTCACGAGCGTACCCGCTAAACAGGCGAA 360
Db 101 GlnProGlyGlnValAlaIleAlaLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 120
Qy 361 GCCTTCCCGCGCGCTGGACCTGACTGTGCAATAACCGTACATCCAGGCGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Qy 421 TCCAGTCCGATCCGTTATTATTAATCTCTAAATCTGGCGTTGCGCAACTGGATTAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Qy 481 AAGTGACTACGGATCTCCTCAGAGCGGAGGCTCAATTCGCTGACTTTTACCGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Qy 541 CGGCAACCGCGTTCGCGAATCGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheSerGlnLeuAsnLeuCys 200
Qy 601 CTTAAACGTCGAGAACACGAGCAAGTGTTCATTAAACGAGGCAATPACCATCGGAATC 660
Db 201 LeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Qy 661 AAGTGAGCGCGCAATCTCTAATTAACCGTCCGTTAAGCTGAGCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Qy 721 GAGATATTCTCTCGCAACAGCACAGGAATCGCGAGCGGGTGGGAAGGATCACCC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Qy 781 GATTACACAGTCGGAACACCTTCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Qy 841 CGCACCGCAGAGTTGCCGAGCGCGCACCCCGCTTATTTGATTGATCATGCGAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300

RESULT 12
US-10-284-962-14
; Sequence 14, Application US/10284962
; Publication No. US20030206913A1
; GENERAL INFORMATION:
; APPLICANT: Webel, Douglas M.
; APPLICANT: Orr, Donald E.
; TITLE OF INVENTION: PHYTASE-CONTAINING ANIMAL FOOD AND METHOD
; FILE REFERENCE: 834460-71725
; CURRENT APPLICATION NUMBER: US/10/284,962
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/335,303
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-284-962-14

Alignment Scores:
Pred. No.:      1.86e-197      Length:      432
Score:          2221.00      Matches:      428
Percent Similarity: 99.07%      Conservative: 0
Best Local Similarity: 99.07%      Mismatches: 4
Query Match:      94.55%      Indels:      0
DB:              15          Gaps:      0

US-10-034-985-1 (1-1323) x US-10-284-962-14 (1-432)

Qy 1 ATGAACCGATCTTAATCCCATTTTATCTCTCTGATTCGGTTAAACCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Qy 61 TTCGCTCAGATGAGCGGAGCTGAAGCTGGAAGTGTGTGATTCAGTCGTCATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Qy 121 GTGCGTCTCAACCAAGGCCACCAACTGATGCAGGATGTCACCCAGAGCATGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Qy 181 ACCTGGCCGTTAAATCTGGTGTGCTGACACCGCGNGGTGTGAGCTAAATCGCCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyLeu 80
Qy 241 GGACATTACCAACCCGAGCGTCTGGTAGCCGACGGATGCTGGCGAAAGGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Qy 301 CAGTCTGTCAGTCGCGATATTCTGATGTCACGAGCGTACCCGCTAAACAGGCGAA 360
Db 101 GlnProGlyGlnValAlaIleAlaLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 120
Qy 361 GCCTTCCCGCGCGCTGGACCTGACTGTGCAATAACCGTACATCCAGGCGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Qy 421 TCCAGTCCGATCCGTTATTATTAATCTCTAAATCTGGCGTTGCGCAACTGGATTAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Qy 481 AAGTGACTACGGATCTCCTCAGAGCGGAGGCTCAATTCGCTGACTTTTACCGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Qy 541 CGGCAACCGCGTTCGCGAATCGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheSerGlnLeuAsnLeuCys 200
Qy 601 CTTAAACGTCGAGAACACGAGCAAGTGTTCATTAAACGAGGCAATPACCATCGGAATC 660
Db 201 LeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Qy 661 AAGTGAGCGCGCAATCTCTAATTAACCGTCCGTTAAGCTGAGCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Qy 721 GAGATATTCTCTCGCAACAGCACAGGAATCGCGAGCGGGTGGGAAGGATCACCC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Qy 781 GATTACACAGTCGGAACACCTTCTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Qy 841 CGCACCGCAGAGTTGCCGAGCGCGCACCCCGCTTATTTGATTGATCATGCGAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
```

```
Db 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
QY 241 GGCATTACCAAGCCAGCGTCTGGTAGCCAGAGATTGTCGGGAAAAAGGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGGTCAAGTCGGGATTATTCGTGATGTCGAGAGGTACCGTAAACAGGCGAA 360
Db 101 GlnProGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCCCGCGGTGGCACCTGACTGTGCAATAACCGGTACATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGATCGGTTATTATCTCTAAACCTGCTAAACCTGGCGTTGGCCAACTGGATAACGGC 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AAGTGTACTCAGCGATCCTCAGCAGGAGGAGGTCAATTGCTGACTTTACCGGSCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGCACAACCGCGTTTCGCAACTGGAAACGGGTGCTTAATTTTCGCAATCAAACTTTGTC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheSerGlnLeuAsnLeuCys 200
QY 601 CTTAAACGTGAGAAACAGGAGAAAGCTGTTCATTACGAGGAGCATTTACCATCGGAATC 660
Db 201 LeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTGTAGCCGACATGCTCATTAACCGGTGCGGTAGCTCGCATCAATCTGTCAGC 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTCTCTGCAACAGCACAGGGAATCCGAGCGCGGGTGGGGAAGATCAC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
QY 781 GATTACACAGTGGAAACCTTGCTAAGTTTGCAATACCGGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
QY 841 CGCAGCGCAGAGGTGCGCGAGCGGCCACCCCGTTATTGATTTGATTCATCGGACGC 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
QY 901 TTGACGCCCATCCACCGCAAAACAGCGGTATGTTGATTCATTCACCTTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTCGCGACAGATACTATCTGGCAATCTCGCGCGCGCACTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTCCTCGGTAGCCGATACACCGCGCGAGTGTGTAACCTGGTGTGAAACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTGGGTAAACGATACAGCCAGTGGATTGAGTTTCGTGCTGTTCCAGACTTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATGGGTGATAAAGCGCGTCTCATTAATAACCGCGCGAGAGGTGAACACTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGCGAGGATGTGAAGAGCGAAATGCGCAGGATGTTGTTGCTGGCAGGTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATGAGCACCGATACCGGCGTGCAGTTTG 1296
```

```
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
RESULT 13
US-10-266-041-1
; Sequence 1, Application US/10266041
; Publication No. US2003007284A1
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2791
; CURRENT APPLICATION NUMBER: US/10/266,041
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US/09/540,149
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 1
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (433)
; OTHER INFORMATION: Xaa at position 433 in this sequence is unknown
US-10-266-041-1
Alignment Scores:
Pred. No.: 1,86e-197 Length: 433
Score: 2221.00 Matches: 428
Percent Similarity: 99.07% Conservative: 0
Best Local Similarity: 99.07% Mismatches: 4
Query Match: 94.55% Indels: 0
DB: 14 Caps: 0
```

US-10-034-985-1 (1-1323) x US-10-266-041-1 (1-433)

```
QY 1 ATCAAAAGCATCTTAATCCCATTTTATCTCTGATTCGTTAACCCCGCAATCTGCA 60
Db 1 MeLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTCAGCGGAGCTGAAGCTGAAAGTGTGGTGAATGTGCTCATGTG 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTCGTCGCTTCAACCAAGCCAGCCAGCACTGATCAGGATGTCACCCAGCGCATGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGGCGGTAAACCTGGGTTGGCTGACACCGCGNGGTGGTGAAGTAAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
QY 241 GGCATTACCAAGCCAGCGTCTGGTAGCCAGAGATTGTCGGGAAAAAGGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGGTCAAGTCGGGATTATTCGTGATGTCGAGAGGTACCGTAAACAGGCGAA 360
Db 101 GlnProGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCCCGCGGTGGCACCTGACTGTGCAATAACCGGTACATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGATCGGTTATTATCTCTAAACCTGCTAAACCTGGCGTTGGCCAACTGGATAACGGC 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AAGTGTACTCAGCGATCCTCAGCAGGAGGAGGTCAATTGCTGACTTTACCGGSCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
```

```
QY 541 CGGCAACGCGGTTTCGGAACCTGGAACGGGTCTTAATTTTCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheSerGlnLeuAsnLeuCys 200
QY 601 CTTAAACGTGAGAACAGGACGAGCTGTTTCATTAAACGAGGCTTACCATCGGAATC 660
Db 201 LeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTGAGCGCCGACAAATGCTCATTAACCGGTGCGGTAAAGCCCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTTCTTCGCAACAGCACAGGAATGCCGAGCGCGGGTGGGAAGATCACC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
QY 781 GATTACACCAAGTGAACACTTCTAAGTTTGATACCGGCATCAATTTTCTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
QY 841 CGCACGCGCAGAGTTGCCGCGCGCCACCGCTTATTGGATTTGATCGGACGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
QY 901 TTGACGCCCCATCCACCGCAAAACAGCGGTATGGTGTGATGACATPACCCACTTCA 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
QY 961 TTATATCCCGACGACGATCTAATCTGGCAATCTCGCGCGCGCTGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTCTCCGCTGAGCGGATAACACCGCCGCGTGTGAGTGTGAGTGTGAGTGTGAG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTGGGTAAAGCGATAACAGCGAGTGTGATTCAGGTTTCGCTGCTTCCAGACTTTAC 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATGGGTATAAAACCGCGCTGCTCATTAATACGCGCGCGCGAGAGGTGAACACTG 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGCGAGATGTGAAGAGCGAAATGCGCAGGCGATGTTGTTGCGCAGGTTTTCAGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATGAAGCAGCATACCGCGCTGCAGTTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
```

RESULT 14

```
US-10-284-962-5
; Sequence 5, Application US/10284962
; Publication No. US20030206913A1
; GENERAL INFORMATION:
; APPLICANT: Weibel, Douglas M.
; APPLICANT: Orr, Donald E.
; APPLICANT: Ruch, Frank E.
; TITLE OF INVENTION: PHYTASE-CONTAINING ANIMAL FOOD AND METHOD
; FILE REFERENCE: 83460-71725
; CURRENT APPLICATION NUMBER: US/10/284,962
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/335,303
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 432
; TYPE: PRN
; ORGANISM: Escherichia coli
```

US-10-284-962-5

Alignment Scores:

```
Pred. No.: 3 18e-195 Length: 432
Score: 2197.00 Matches: 424
Percent Similarity: 98.61% Conservative: 2
Best Local Similarity: 98.15% Mismatches: 6
Query Match: 93.53% Indels: 0
DB: 15 Gaps: 0
```

US-10-034-985-1 (1-1323) x US-10-284-962-5 (1-432)

```
QY 1 ATGAAGAGCATCTAATCCCATTTTATCTCTTCTGATTCGTTAAACCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTCAGCGGAGCTGAAGCTGGAAGTGTGTGATTGTTCAGTCGTCATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTGCTGCTCCAAACGAGCCAGCAACTGATGCAGGATGTCACCCAGACGCGATGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGCGCGTAAACCTGCGTGTGCTGACACCGCGGNGTGTGAGCTAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
QY 241 GGACATTACCAACGCCAGCGCTGTTAGCCGACGGATTGCTGGCGAAAGGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGCTGTGTCAGTCGCGGATTATTGCTGATGTGTCGAGCGGTACCCGCTAAACAGGCGAA 360
Db 101 GlnProGlyGlnValAlaIleIleValAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCGCGCGGTGCGACCTGACTGTGCAATAACCTACATACCCAGGCGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCCAGTCCCGATCCGTTATTTAATCTCTAAACCTGGCGTGTGCAACTGGATAACGCG 480
Db 141 SerSerProAspProLeuPheIleProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AAGCTGACTCAGCGCATCTCTCAGCAGGCGAGGAGGTCAATTTGCTGCTTTTACCGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAACGCGGTTTCGGAACCTGGAACCGGTGCTTAATTTTCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuAsn 200
QY 601 CTTAAACGTGAGAACAGGACGAGCTGTTTCATTAAACGCGGTGCGGTAAAGCCCTCGCAT 660
Db 201 LeuLysArgGluLysGlnAsnGluSerCysAsnLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTGAGCGCCGACAAATGCTCATTAACCGGTGCGGTAAAGCCCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTTCTTCGCAACAGCACAGGAATGCCGAGCGCGGGTGGGAAGATCACC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
QY 781 GATTACACCAAGTGAACACTTCTAAGTTTGATACCGGCATCAATTTTCTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
QY 841 CGCACGCGCAGAGTTGCCGCGCGCCACCGCTTATTGGATTTGATCGGACGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTGACGCCCCATCCACCGCAAAACAGCGGTATGGTGTGATGACATPACCCACTTCA 960
```

Db 301 LeuThrProHisProGlnLysGlnAlaTyGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTGCCGACGACATATACTGGCAAAATCTCGCGCGGCACCTGAGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACCTTCCCGGTGAGCGGATACACGCCCCAGGTGGTGAACCTGTTTGAACCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGluArgTrp 360
QY 1081 CGTCGGCTAAGCAGTAAACGCCAGTGGATTTCAGTTTCGGTGTCTTCCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATCGCTGATAAAACGCCGCTGTCAATTAATAACGCCGCCCGGAGAGGTGAACCTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGGCAGGATGTAGAGCGAAATGCCAGGCGATGTTTCGTTGGCAGGTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCTGTAATGAAGCAGCATACCGCGCTGCAGTTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

RESULT 15

US-10-156-660-2
; Sequence 2, Application US/10156660
; Publication No. US20030103958A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson R.
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM
; TITLE OF INVENTION: AND METHODS FOR MAKING AND USING THEM
; FILE REFERENCE: 09010-029007
; CURRENT APPLICATION NUMBER: US/10/156,660
; PRIOR FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 436

TYPE: PRT

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: modified phytase enzyme

US-10-156-660-2

Alignment Scores:

Pred. No.: 2,19e-194 Length: 436
Score: 2188.00 Matches: 425
Percent Similarity: 97.70% Conservative: 0
Best Local Similarity: 97.70% Mismatches: 10
Query Match: 93.15% Indels: 0
DB: 14 Gaps: 0

US-10-034-985-1 (1-1323) x US-10-156-660-2 (1-436)
QY 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTCTGATTTCGGTTAAACCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTGAGCCGGAGCTGAAGCTGGAAAGTGTGGTATTGTCTAGTCTCATGCT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTGCGTGTCTCAACCAAGGCCACGCAATGATGCGAGATGTCAACCCGACGCGATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACTGCGCCGTTAAACTGGTGTGGCTGCACACCGCGNGGTGGTGAGCTAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyGluLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
QY 241 GGACATTACCAACGCCGAGCGTCTGTAGCCGCGAGATTGCTGGCGAAAAGGGCTGCCCG 300
Db 81 GlyHisTyTrpArgGlnArgLeuValAlaAspGlyLeuLeuProLysCysGlyCysPro 100
QY 301 CAGTCTGTGTCAGGTCCGATTATTGCTGATGTCGACGAGCGTACCCTGTAACACGCGAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCCCGCGGTGGCGACCTGACTGTGCAATAACCGTACATACCCAGGCGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCCAGTCCCGATCCGTTATTATCTCTAAAAACCTGGCGTTTGCACACTGGATAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AACGTGACTGACCGGATCTCAGACGGGAGGAGGTCAATTGCTCAGCTTTACCGGGCAT 540
Db 161 AsnValThrAspAlaIleLeuGluArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAAAACGGCGTTTCGCAACCTGGAACCGGTCTTAATTTCCGCAATCAAACTGTGTC 600
Db 181 TyrGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTGAGAAACACGAGCAGAAAGCTGTTCAATTAACGAGCGCATTACCATCGAACTC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTGAGCGCCGACAAATGCTCTCAATTAACCGGTGCGGTAAAGCCCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspCysValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTTCTCTGCAACAGCACAGGGAATCCCGAGCGCGGGTGGGGAAGGATCACC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTTCACCAAGTGAACACCTTGTGTAAGTTTGCATAACCGCAATTTTATTGCTGTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 280
QY 841 CGCACGCCAGAGTTGCCCGCAGCGCCGCCCGTATTGATTTGATTCATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTGACGCCCATCCACCGCAAAAACAGGCGGTATGGTGTGACATTACCCTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTCCCGGACGACATATACTGCGCAAAATCTCGCGCGCGCATGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTCGCCGTCAGCCGGATAACACGCCCGAGGTGGTGAACCTGTTTGAACGCTGG 1080

Db	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTyr	360
Qy	1081	CGTCGGCTAAGCGATAACAGCCAGTGGATTTCAGGTTTCGCTGGTCTCTCCAGACTTTACAG	1140
Db	361	ArgArgLeuSerAspAsnSerGlnTyrIleGlnValSerLeuValPheGlnThrLeuGln	380
Qy	1141	CAGATGGCTGATAAAACGCCGCTGTCAATTAATACGCCGCCGCCGAGAGGTCAAACTCACC	1200
Db	381	GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr	400
Qy	1201	CTGCGCAGATGTCGAAGACGAAATGCCACGGCATGTGTTTCGTTGGCAGGTTTTACGCCAA	1260
Db	401	LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln	420
Qy	1261	ATCTGTAATGAGCAGCCGATACCGCGCTGCAGTTTGGATCTCAT	1305
Db	421	IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHis	435

Search completed: March 12, 2005, 00:05:49
Job time : 209.5 secs

THIS PAGE BLANK (user #)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 11, 2005, 23:20:13 ; Search time 45 Seconds
(without alignments)
4389.365 Million cell updates/sec

Title: US-10-034-985-1
Perfect score: 2349
Sequence: 1 atgaagcgatcttaatccc.....atcacatcaccatcactaa 1323

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 1027090

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10034985/runat 10032005_164241_11973/app_query.fasta_1.1479
-DB=issued Patents AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10034985 @CGN 1.1 46 @runat 10032005_164241_11973 -NCPUS=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFBLCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2302	98.0	440	3	US-09-259-214-2
2	2302	98.0	440	3	US-09-318-528-2
3	2302	98.0	440	3	US-09-231-931-2
4	2302	98.0	440	4	US-09-580-515-2
5	2221	94.6	433	4	US-09-540-149A-1
6	2194.5	93.4	423	2	US-08-910-798-2
7	528	22.5	421	4	US-09-489-039A-7512
8	376	16.0	522	4	US-09-489-039A-13501
9	119	5.1	439	3	US-09-044-718-3
10	119	5.1	439	4	US-10-062-848-3
11	119	5.1	440	4	US-09-684-855-106
12	119	5.1	440	4	US-09-684-855-128

13	119	5.1	440	4	US-09-684-855-151	Sequence 151, Appl	
14	119	5.1	440	4	US-09-488-265B-6	Sequence 6, Appli	
15	119	5.1	449	3	US-09-044-718-12	Sequence 12, Appl	
16	119	5.1	449	3	US-10-062-848-12	Sequence 12, Appl	
17	119	5.1	465	3	US-08-868-435-33	Sequence 33, Appl	
18	119	5.1	465	3	US-08-744-231-33	Sequence 33, Appl	
19	119	5.1	465	3	US-09-044-718-78	Sequence 78, Appl	
20	119	5.1	465	4	US-09-636-499-6	Sequence 6, Appli	
21	119	5.1	465	4	US-09-273-871A-8	Sequence 8, Appli	
22	119	5.1	465	4	US-10-083-452-8	Sequence 8, Appli	
23	119	5.1	465	4	US-09-635-504-33	Sequence 33, Appl	
24	119	5.1	465	4	US-10-062-848-78	Sequence 78, Appl	
C	25	118.5	5.1	575	4	US-09-252-991A-20649	Sequence 20649, App
26	118	5.0	440	4	US-09-684-855-107	Sequence 107, App	
27	118	5.0	440	4	US-09-684-855-129	Sequence 129, App	
28	118	5.0	440	4	US-09-684-855-152	Sequence 152, App	
29	118	5.0	440	4	US-09-488-265B-7	Sequence 7, Appli	
30	117	5.0	386	1	US-08-758-213-1	Sequence 1, Appli	
31	117	5.0	386	2	US-08-692-787-48	Sequence 48, Appli	
32	117	5.0	386	3	US-09-097-199-48	Sequence 48, Appl	
33	117	5.0	386	4	US-09-949-016-6022	Sequence 6022, Ap	
34	117	5.0	515	2	US-09-146-283-2	Sequence 2, Appli	
35	117	5.0	515	3	US-08-579-823A-2	Sequence 2, Appli	
36	117	5.0	515	3	US-09-344-195-2	Sequence 2, Appli	
C	37	116.5	5.0	330	4	US-09-252-991A-32186	Sequence 32186, A
38	116.5	5.0	442	4	US-09-949-016-9351	Sequence 9351, Ap	
39	116.5	5.0	442	4	US-09-949-016-9352	Sequence 9352, Ap	
40	116.5	5.0	442	4	US-09-949-016-9353	Sequence 9353, Ap	
41	116.5	5.0	442	4	US-09-949-016-9354	Sequence 9354, Ap	
42	116.5	5.0	442	4	US-09-949-016-9355	Sequence 9355, Ap	
43	116.5	5.0	450	4	US-09-765-111A-14	Sequence 14, Appl	
C	44	116.5	5.0	1093	4	US-09-252-991A-21827	Sequence 21827, A
C	45	115.5	5.0	539	4	US-09-252-991A-33061	Sequence 33061, A

ALIGNMENTS

RESULT 1
US-09-259-214-2
; Sequence 2, Application US/09259214A
; Patent No. 6110719

GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: DIVER1370-1
; CURRENT APPLICATION NUMBER: US/09/259,214A
; CURRENT FILING DATE: 1999-03-01
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-259-214-2

Alignment Scores:
Pred. No.: 9.08e-230 Length: 440
Score: 2302.00 Matches: 440
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.00% Indels: 0
DB: 3 Gaps: 0

US-10-034-985-1 (1-1323) x US-09-259-214-2 (1-440)

Qy	1	ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGTTAACCCGCAATCGCA	60
Db	1	MeLysAlaIleuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla	20
Qy	61	TTGCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGATGTTCAGTCGTCTGTT	120


```
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Qy 481 AACGTGACTACCGATCTCTCAGACGGCAGAGGGTCAATGCTGACTTTTACCGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Qy 541 CGGCAACCGCGTTTCGGCACTCGAACGGTGTCTTAATTTCCGCAATCAAACTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Qy 601 CTTAAACGTCAGAAACAGGACGAAAGCTGTTTCAATACGACGAGCATTAACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Qy 661 AAGGTGAGCCGCGCAATGCTCATTAACCGGTGCGTAAAGCTTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Qy 721 GAGATATTTCTCTGCAACAGCACAGGGAATCCGAGCGCGGGTGGGAAGGATCAC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Qy 781 GATTTCACACAGTGGAAACCTTGCTAAGTTTGCTAAACGCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Qy 841 CGCACGCGCAGAGTTGCCGACGCGCCACCCGCTTATTTGATTTGATCATGCGACG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Qy 901 TTGACGCCCCATCCACCGCAAAACAGCGGTATGGTGTGACATTAACCATTCAGTACTG 960
Db 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Qy 961 TTTATTGCGCGACACGATCTAATCTGCAAAATCTCGCGCGCGCACTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
Qy 1021 ACCTTTCCCGGTGAGCGGATAACACGCGCGAGGTGGTGAACCTGGTGTGAACTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Qy 1081 CGTCGGCTAAGCATACACCGCAGTGGATTCAGTTTCGTTTCGTTCTCCAGACTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Qy 1141 CAGATGCGTATAAAACCGCGCTGTCAATTAATAACGCGCGCGAGAGGTGAACTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Qy 1201 CTGCGCAGGATGTGAAGACGAAATGCGCAGGCGATGTTGTTGTTGGCAGGTTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Qy 1261 ATCGTGAATGACGACGATACCGCGGTGAGTTTGAGATCTCATACCATCATCATCAC 1320
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
```

RESULT 3

```
US-09-291-931-2
; Sequence 2, Application US/09291931A
; Patent No. 6190897
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/291,931A
; CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
```

```
; SOFTWARE: FastSEQ for Windows Version 4.0.
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-291-931-2

Alignment Scores:
Pred. No.: 9,08e-230 Length: 440
Score: 2302.00 Matches: 440
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.00% Indels: 0
DB: 3

US-10-034-985-1 (1-1323) x US-09-291-931-2 (1-440)
Qy 1 ATCAAAAGCATCTTAATCCCATTTTATCTCTCTGATTCGGTTAAACCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Qy 61 TTCGCTCAGAGTCAGCGGAGCTGAAGCTGAAAGTGTGTGATTTGTCACTCGTCATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
Qy 121 GTGCGTGTCTCAACCAAGGCCACGCAACTGATGTCAGGATGTCACCCAGACGATGCGCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Qy 181 ACCTGCGCGGTAAACCTGGTGGCTGACACCGCGNGGTGGTGAGCTAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaIle 80
Qy 241 GGACATTAACCAACCGCAGCGTCTGTAGCCGCGAGTTGTGGCGAAAAAGGGCTGCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Qy 301 CAGTCTGGTCAGTCCGATATTGTCGTGATGTCGACGAGCGTACCCGTAAAAACAGCGAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Qy 361 GCCTTCGCGCGCGGTGGCCACTGCTGCAATACCGTACATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Qy 421 TCCAGTCCCGATCCGTTATTATCTCTAAAACTGGCGTTTCCCACTGGATAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Qy 481 AACGTCACTGACCGCATCTCTACGAGGCGAGGAGGTCAATTGCTGATTTACCGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Qy 541 CGGCAACCGCGTTTCGGCACTCGAACGGTGTCTTAATTTCCGCAATCAAACTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Qy 601 CTTAAACGTCAGAAACAGGACGAAAGCTGTTTCAATACGCGAGCATTAACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Qy 661 AAGGTGAGCGCGCAATGCTCATTAACCGGTGCGTAAAGCTTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Qy 721 GAGATATTTCTCTGCAACAGCACAGGGAATCCGAGCGCGGGTGGGAAGGATCAC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Qy 781 GATTTCACACAGTGGAAACCTTGCTAAGTTTGCTAAACGCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
```

```
QY 841 CGCAGCCGAGAGTGTCCCGCAGCGCCGACCCCGTATTGATTTGATCATGGCAGCG 900
Db ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuMetAlaAla 300
QY 901 TTGACGCCCATCCACCGCAAAACAGCGGTATGTTGTGATTAACCATTCAGTACTG 960
Db LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTGCCGGACAGCATACTTAATCTCGCAAAATCTCGCGCGCGACCTGGAGCTCAACTGG 1020
Db PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTCCCGGTGAGCGGATTAACACCGCCCGAGTGTGAATCTGTTTGAACGCTGG 1080
Db ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTCGGGTAAAGCATTAACAGCCAGTGGATTTCAGTTTCGGTGTCTTCAGACTTTACAG 1140
Db ArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATGGTGATAAAACGCGCTGTCTTAATAATACGCGCGCGAGAGGTGAAACTGACC 1200
Db GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGCGAGGATGTGAAGAGCGAAATGCGCAGCGCATGTGTTCTGTTGCGAGGTTTTACGCAA 1260
Db LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATGCTGAATGAAGCAGCATACCGGCGTGAGTTTGAGATCTCATCACCATCACCATCAC 1320
Db IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis 440
```

RESULT 4

```
US-09-580-515-2
; Sequence 2, Application US/09580515
; Patent No. 6720014
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/580,515
; CURRENT FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-580-515-2
```

```
Alignment Scores:
Pred. No.: 9,08e-230 Length: 440
Score: 2302.00 Matches: 440
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.00% Indels: 0
DB: 4 Gaps: 0
```

US-10-034-985-1 (1-1323) x US-09-580-515-2 (1-440)

```
QY 1 ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGGTTAAACCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTAGCGGAGCTGAAGCTGGAAAGTGTGTGATTGTCAGTCGTATGGT 120
```

```
Db PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTGCGTGTCCACACCGCCAGCCCAACTGATCAGAGATGTACCCAGACGATCGCCA 180
Db ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGCGCGTAAACCTGGGTTGGCTTGACACCCCGNGGTGGTGAAGCTTAATCGCTATCTC 240
Db ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
QY 241 GGCATTATCAACCCAGCGGTCTGGTAGCGCAGGATTCTCTGCGCAAAAAGGCTGCCCG 300
Db GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGTCAGTCCGCGATTATTGCTGATGTGACGAGCGGTACCGGTAAACAGCGGAA 360
Db GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCGCGCGGTGCGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG 420
Db AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGATCCGTTATTATCTCTAAAACTGGCGTTTGGCCAACTGGATACGCG 480
Db SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AACGTCAGTCAGCGATCTCTCAGCAGGCGAGGAGGTCAATTCTGACTTTTACGGGCAT 540
Db AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGCAAAACGCGGTTTCGCAACTGGAAACGCGTCTTAAATTTTCGCAATCAAACTTTGTGC 600
Db ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTGAGAAAACAGACGAAAGCTGTTTCAATTAAACGAGCATTAACATCGGAATC 660
Db LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTCAGCGCGCACAAATGTCATTAAACGCGTCCGTAAGCCTCCATCAATCAATCTGACG 720
Db LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTTCTCTCGCAACAAGCACAGGGAATCCCGAGCGCGGTGGGGAAGATCAC 780
Db GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTCAACAGTGGAAACCTTGTGTAAGTTTGCATAACGCGCAATTTTATTTGCTACAA 840
Db AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
QY 841 CGCACGCCAGAGGTTCCCGCAGCGCCCGCTTATTTGATTGATTCATGCGCAGCG 900
Db ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
QY 901 TTGACGCCCATCCACCGCAAAACAGCGGTATGTTGTGACATTAACCATTCAGTACTG 960
Db LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTGCCGGACAGCATACTTAATCTCGCAAAATCTCGCGCGCGACCTGGAGCTCAACTGG 1020
Db PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTCCCGGTGAGCGGATTAACACCGCCCGAGTGTGAATCTGTTTGAACGCTGG 1080
Db ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTCGGGTAAAGCATTAACAGCCAGTGGATTTCAGTTTCGGTGTCTTCAGACTTTACAG 1140
Db ArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATGGTGATAAAACGCGCTGTCTTAATAATACGCGCGCGAGAGGTGAAACTGACC 1200
```

Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Qy 1201 CTGCAGGATGTGAAGAGCGAAATGCCAGGGCATGTGTTGGTGGCAGGTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Qy 1261 ATCGTGATGACGACGACCATACCGCGGTGAGTTTGAGATCTCATCACCACCATCAC 1320
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis 440

RESULT 5
US-09-540-149A-1
; Sequence 1, Application US/09540149A.
; Patent No. 6511699
; GENERAL INFORMATION:
; APPLICANT: Lei, Kingen
; TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2791
; CURRENT APPLICATION NUMBER: US/09/540,149A
; CURRENT FILING DATE: 2000-03-31
; PRIOR FILING DATE: 60/127,032
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (433)
; OTHER INFORMATION: Xaa at position 433 in this sequence is unknown
US-09-540-149A-1

Alignment Scores:

Prod. No.: 2,26e-221 Length: 433
Score: 2221.00 Matches: 428
Percent Similarity: 99.07% Conservative: 0
Best Local Similarity: 99.07% Mismatches: 4
Query Match: 94.55% Indels: 0
DB: 4 Gaps: 0

US-10-034-985-1 (1-1323) x US-09-540-149A-1 (1-433)

Qy 1 ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGGTTAAACCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Qy 61 TTCGCTCAGAGTACGCGGAGCTGAAGCTGGAAGTGTGTGATTCAGTCGTCTATGTT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Qy 121 GTGGTGCTCCAAACGAGCGCAGCACTGATGCGAGGATCTACCCGACGAGCGGCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Qy 181 ACCTGGCGGTAAACCTGGGTGTGCTGACACCGCGGTGTGAGCTAACTCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
Qy 241 GGACATTAACCAACCGCGCTGTGTTAGCCGACGATTCGCGGAAAAAGGCGTCCCG 300
Db 81 GlyHisThrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Qy 301 CAGTCGTGTCAGTCGCGATTTATGCTGATGTCGACGAGGTACCGTAAACAGCGGAA 360
Db 101 GlnProGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Qy 361 GCCTTCGCGCGCGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Qy 421 TCCAGTCCCGATCCGTTATTTAATCTCTTAAAAAATCGCGGTTCGCCAACTGGATAACGCG 480

Db 141 SerSerProAspProPhePheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Qy 481 AAGCTGACTGACCGCATCTCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTTACGGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Qy 541 CGGCAAAACGGCGTTTCGCAACTGGAAACGGGTGCTTAATTTCCGCAATCAAACTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheSerGlnLeuAsnLeuCys 200
Qy 601 CTTTAAACGTGAGAAACAGGACGAAAGCTGTTCTAATACCGCAGGACATTAACCATCGAACTC 660
Db 201 LeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Qy 661 AAGGTGAGCGCCGACCAATGTCTCATTAAACCGGTGCGGTAAAGCCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Qy 721 GAGATATTCTCTCTCAACAGCAGGGAATCCGAGCGCGGGTGGGGAAGATCAC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Qy 781 GATTTCACACGAGTGAACACCTTGTAACTTGCATAACGCGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
Qy 841 CGCAGCGCAGAGGTTCGCCGACGCGCCACCGCTTATTGATTTGATCATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Qy 901 TTGACGCCCATCCACCGCAAAACAGGCGTATGGTGTGATTAATCCCATTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyxGlyValThrLeuProThrSerValLeu 320
Qy 961 TTTATTGCGGACACGATCTAATCTGGCAAAATCTCGCGCGCCTGAGGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
Qy 1021 ACCTTTCGCGTCCGCGGATAACCGCGCCAGGCGGTGTAACCTGGTGTGTTGAACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Qy 1081 CGTCGCTAAGCGATAACACGCGATGATTCAGGTTTCGCTGCTTCCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Qy 1141 CAGATCGGTGATAAAACCGCGCTGCTATTAAATACGCCGCCGCGAGAGGTGAACGAC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Qy 1201 CTGCGAGGATGTGAAGAGCGAAATGCGAGGCGATGTGTTGCTGGCAGGTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Qy 1261 ATCGTGAATGAACGACGACATACCGCGGTGCAGTTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

RESULT 6

US-08-910-798-2
; Sequence 2, Application US/08910798
; Patent No. 5876997
; GENERAL INFORMATION:
; APPLICANT: KRETZ
; TITLE OF INVENTION: NOVEL PHYTASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: US

```

, ZIP: 92037
,
, COMPUTER READABLE FORM:
,
, MEDIUM TYPE: Floppy disk
,
, COMPUTER: IBM PC compatible
,
, OPERATING SYSTEM: PC-DOS/MS-DOS
,
, SOFTWARE: PatentIn Release #1.0, Version #1.25
,
, CURRENT APPLICATION DATA:
,
, APPLICATION NUMBER: US/08/910,798
,
, FILING DATE: August 13,1997
,
, CLASSIFICATION: 435
,
, ATTORNEY/AGENT INFORMATION:
,
, NAME: HAILE, PH. D., LISA A.
,
, REGISTRATION NUMBER: 38,347
,
, REFERENCE/DOCKET NUMBER: 09010/029001
,
, TELECOMMUNICATION INFORMATION:
,
, TELEPHONE: 619/678-5070
,
, TELEFAX: 619/678-5099
,
, INFORMATION FOR SEQ ID NO: 2:
,
, SEQUENCE CHARACTERISTICS:
,
, LENGTH: 423 amino acids
,
, TYPE: amino acid
,
, TOPOLOGY: linear
,
, MOLECULE TYPE: protein
,
, US-08-910-798-2

```

164	ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeu	Cys	183
601	CTTAAACGTCAGAAACAGGACGAAAGCTGTTCTATTAAACGACGACCATTTACCATCGGAATC		660
184	LeuIysArgGluIysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu		203
661	AAGGTGAGCCGCACAAATGTCTCATTAACCGGTGCGGTAAAGCTTCGCATCAATGCTGACG		720
204	LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr		223
721	GAGATATTTCTCTGCACAAAGACACAGGNAATCCCGACCGCGGGTGGGGAAGGATCACC		780
224	GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr		243
781	GATTTCACACAGTGGAAACACTTGTCTAAGTTTCGATAACGCGCAATTTTATTGCTTACAA		840
244	AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln		263
841	CGCACGCACAGGTTCGCCGACCGCGCCACCCCGTTATTGGATTGTATCATGGCAGCG		900
264	ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla		283
901	TTGACGCCCATCCACCGGCAAAACAGGGGTATGGTGTGACATTACCCACTTCAGTACTG		960
284	LeuThrProHisProGlnIysGlnAlaTyrGlyValThrLeuProThrSerValLeu		303
961	TTTATTCCGGACACCATACTAATCTGGCAAACTCGGCGGCGCATCTGAGAGCTCAACTGG		1020
304	PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp		323
1021	ACGTTTCCCGGTACAGCCGGATAACGCGCCAGGTGGTGAACTGGTGTGTTGAACGCTGG		1080
324	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp		343
1081	CGTCGGCTAAGCGATAACAGCCAGTCGATTACGTTTCGCTGGTCTTCAGACTTTACAG		1140
344	ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln		363
1141	CAGATCGGTGATAAACCGCGTGTCTATTAAATACCGCGCGGAGAGGTGAACATGACC		1200
364	GlnMetArgAspIysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr		383
1201	CTGGCAGGATGTCAAGAGCGAAATGCGAGGGCATGTGTTGTTGGCAGGTTTACGCAA		1260
384	LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln		403
1261	ATCGTGAATGAACGACGATACCGCGGTGCGAGTTTCAGATCTCATCACCATCACCATCAC		1320
404	IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHisHis		423

Db 324 AlaSerLeuLeuThralaLeuAapPheLysProTyrClnLeuProGlyGlnTyrGluArg 343
 QY 1045 ACGCCCGCAGGTGGAACACTGGTTGTGAACGCTGGCTGGCTAAACGACGACG 1104
 Db 344 ThrProileGlyLysLeuLeuPheGlnArgTrpHisAspSerAlaGlyAsnArgAsp 363
 QY 1105 TGAATTCAGGTTTCGCTGGTCTTCCAGACTTTACAGCAGATGCGTGATAAAACCCCGCTG 1164
 Db 364 LeuMetLysIleGluTyrValTyrGlnSerThrGluGlnLeuArgAsnAlaAspAlaLeu 383
 QY 1165 TCATTAATAACGCCCGCAGAGAGGTGAACACTGCACCTGCGCAGATGCGAAGACGCGAAAT 1224
 Db 384 ThrLeuGlnAlaProProGlnArgValThrLeuAlaLeuAsnGlyCysProVal--Asp 402
 QY 1225 GCGCAGGCGATGTTTCGTTGGCAGGTTTTCACCAATCTGTAATGAAGCA 1275
 Db 403 AspGlnGlyPheCysProLeuGluThrPheLysMetValIleAsnGluAla 419
 RESULT 8
 US-09-489-039A-13501
 ; Sequence 13501, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBS
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489, 039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 13501
 ; LENGTH: 522
 ; TYPE: PRT
 ; ORGANISM: Klebsiella pneumoniae
 US-09-489-039A-13501
 Alignment Scores:
 Pred. No.: 5,08e-30 Length: 522
 Score: 376.00 Matches: 115
 Percent Similarity: 42.70% Conservative: 75
 Best Local Similarity: 25.84% Mismatches: 209
 Query Match: 16.01% Indels: 46
 DB: 4 Gaps: 15
 US-10-034-985-1 (1-1323) x US-09-489-039A-13501 (1-522)
 QY 4 AAGCGATCTTAATCCACTTTTATCTCTCTGATTCGGTTAACCCCG---CAATCTGCA 60
 Db 107 GlnGlyLeuLeuArgLeuPheIleAlaCysAlaLeuProLeuAlaLeuGlnSerAla 126
 QY 61 TTCGCTCAGAGTCAGCCGAGCTGAAGCTGGAAGTGTGTGATGTCAGTCGTCATGGT 120
 Db 127 AlaAla-----AlaAspTrpGlnLeuGluLysValValGluLeuSerArgHisGly 143
 QY 121 GTCGGTCTCCAAACCAAGGCCAGC---CAACTGATGCGAGGATGTCACCCAGCGCATGG 177
 Db 144 IleArgProThrAlaGlyAsnArgGluAlaIleGluAlaAlaThrGlyArgProTrp 163
 QY 178 CCAACTCTGCGCGTAAACTGGTGTGCTGACACCGCGNGGTGTGTAGCTTAATCGCTAT 237
 Db 164 ThrGluTrpThrHisAspGlyGluLeuThrGlyHisGlyTyrAlaAlaValAsn 183
 QY 238 CTCGGACATTAACACGCCAGCGCTCTGGTAGCCGAGCTGTCGGCGAAAAAGGCTGC 297
 Db 184 LysGlyArgAlaGluGlyGlnHisTyrArgGlnLeuGlyLeuLeu---GlnAlaGlyCys 202
 QY 298 CCCAGCTTCGTCAGGTCGCGATTATTGCTGATGTCGACGACGCTACCCGTAAAAACGGC 357
 Db 203 ProThrAlaGluSerIleTyrValArgAlaSerProLeuGlnArgThrArgAlaThra 222
 QY 358 GAAGCCTTCGCCCGCGGCTGGCACCTGACTGTGTCATAAACCCTGACATACCCAGCAGAT 417

```
Db 223 GlnAlaLeuValAspGlyAlaPheProGlyCysGlyValAlaIleHisTyrVal----- 240
Qy 418 ACGTCCAGTCCGATCCGTTATTTAATCTCTAAACTCGGTTCCCAACTGGATAAC 477
Db 241 SerGlyAspAlaAspProLeuPheGlnThrAspLysPheAlaAlaThrGlnThrAspPro 260
Qy 478 GCGAACGTGACTGACGGATCCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTACCGGG 537
Db 261 AlaArgGlnLeuAlaAlaValLysGluLysAlaGly-----AspLeuAlaGln 276
Qy 538 CATCGCAACCGCGTTTCGCAACTGGAACGGGTCTTAATTTTCGCAATCAAACTTG 597
Db 277 ArgArgGlnAla-----LeuAlaProThrIleGlnLeuLeuLysGlnAlaVal 292
Qy 598 TGCCTTAACGTGAGAAACAGGACGAAAGCTGTTCATTAACGAGGATTAACATCGGAA 657
Db 293 Cys-----GlnAlaAspLysProCysProIlePheAspThr---ProTrpGln 307
Qy 658 CTCAGGTGAGCGCGACATGCTCATTAACCGGTGCGGTAGCCTCGCATCAATGCTG 717
Db 308 ValGluGlnSerLysSerGlyLysThrIleSerGlyLeuSerValMetAlaAsnMet 327
Qy 718 ACGAGATATTTCTCTCGCAACAGCAGCGAATGCCG-----GAGCGGGGTGGGA 771
Db 328 ValGluThrLeuArgLeuGlyTrpSerGluAsnLeuProLeuSerGlnLeuAlaTrpGly 347
Qy 772 AGGATACCGATTCACACAGTGGAAACCTTGCTGAAGTTTGATTAACGCGCAATTTAT 831
Db 348 LysIleThrGlnAlaArgGlnIleThrAlaLeuLeuProLeuLeuThrGluAsnTyrAsp 367
Qy 832 TTGCTACACGCGCAGAGGTGCGCGAGCGCGCCACCCGCTATTGATTTGATC 891
Db 368 LeuThrAsnAspValLeuTyrThrAlaGlnLysArgGlySerValLeuLeuAsnAlaMet 387
Qy 892 ATGCGACGTGACGCGCCATCCACCGCAAAACAGCGGTATGCTGTGACATTACCCACT 951
Db 388 LeuAspGlyValLysProGluAlaAsnProAsnValArgTrp----- 401
Qy 952 TCAGTACTGTTATTCGCGACAGCATACTAATCTGGCAATCTCGCGCGGCACTGGAG 1011
Db 402 -----LeuLeuLeuValAlaHisAspThrAsnIleAlaMetValArgThrLeuMetAsn 419
Qy 1012 CTCAACTGGACGCTTCCCGT---CAGCGGATAACACGCGCGAGTGTGTAAGTGGTG 1068
Db 420 PheSerTrpGlnLeuProGlyTyrSerArgGlyAsnIleProProGlySerSerLeuVal 439
Qy 1069 TTGAACTGCGGTGCGGTAAAGCGATAACAGCGAGTGGATTTCAGGTTTCGCTGCTTC 1128
Db 440 LeuGluArgTrpArgAsnAlaLysSerGlyGluArgTyrLeuArgValTyrPheGlnAla 459
Qy 1129 CAGACTTACAGCATGCGTGATAAAACCGCGCTGTCTAATTAATACGCGCGCGGAGAG 1188
Db 460 GlnGlyLeuAspAspLeuArg-----ArgLeuGlnThrProAspAlaGln 474
Qy 1189 GTGAACTGACC-----CTGCGAGGATGTAAGCGCAATGCGCAGGCG 1233
Db 475 HisProMetLeuArgGlnGluTrpHisGlnProGlyCysArgGlnThrAspValGlyThr 494
Qy 1234 ATGTGT-----TCGTTGGCAGGTTTACGCAATCTGTAATGAAGACGCGATA 1281
Db 495 LeuCysProPheGlnAlaAlaIleThrAlaLeuGlyGlnArgIleAspArgSerSerAla 514
Qy 1282 CCGCGTGCAGTTG 1296
Db 515 ProAlaValAlaMet 519
```

RESULT 9

US-09-044-718-3

; Sequence 3, Application US/09044718

; Patent No. 6391605

; GENERAL INFORMATION:

; APPLICANT: KOSTREWA, Dirk

```
; APPLICANT: PASAMONTES, Luis
; APPLICANT: TOMSCHY, Andrea
; APPLICANT: van LOON, Adolphus
; APPLICANT: VOGEL, Kurt
; APPLICANT: WYSS, Markus
; TITLE OF INVENTION: MODIFIED PHYTASES
; FILE REFERENCE: Modified Phytases
; CURRENT APPLICATION NUMBER: US/09/044,718
; CURRENT FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: EP 97810175.6
; PRIOR FILING DATE: 1997-03-25
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-09-044-718-3

Alignment Scores:
Pred. No.: 0.00207 Length: 439
Score: 119.00 Matches: 102
Percent Similarity: 36.67% Conservative: 70
Best Local Similarity: 21.75% Mismatches: 183
Query Match: 5.07% Indels: 114
DB: 3 Gaps: 24
```

US-10-034-985-1 (1-1323) x US-09-044-718-3 (1-439)

```
Qy 19 CCATTTTATCTCTCTTCGATTCCGTTAAACCCCAATCTGCATTCGTCAGAGTGAGCCG 78
Db 26 ProPheSerLeuGluAspGluLeuSerValSerSerLysLeuProLys----- 42
Qy 79 GAGCTGAAGCTGGAAGTGTGTGATTGTCAAGTCGTATGTCGTGTGTGTGTGTCAACC--- 135
Db 43 AspCysArgIleThrLeuValGlnValLeuSerArgHisGlyAlaArgTyrProThrSer 62
Qy 136 -----AAGCCACGCAACTGATGACGAGTGTACCCCGACGACGCA----- 174
Db 63 SerLysSerLysLysTyrLysLysLeuValThrAlaIleGlnAlaAsnAlaThrAspPhe 82
Qy 175 -----TGGCCAACTGCGCGGTAAACTGAGTGGTGG-----CTGACA 210
Db 83 LysGlyLysPheAlaPheLeuLysThrTyrAsnTyrThrLeuGlyAlaAspAspLeuThr 102
Qy 211 CCGCGNGTGTGAGTAAATCGCTATCTCGGA-----CATTACCAACCCAGCGTCTG 264
Db 103 ProPheGlyGluGlnGlnLeuValAsnSerGlyIleLysPheTyrGlnArgTyrLysAla 122
Qy 265 GTAGCCGACGATTGTGTGGCGAAAGGGTGTCCCGCAGTCTGTGTGAGTGTGCGATTATT 324
Db 123 LeuAlaArgSerValPro-----PheIleArg 132
Qy 325 GCTGATGTCGACGAGGTACCCGTAAACAGCGGAGACCTTCGCGCGCGGCTG----- 378
Db 133 AlaSerGlySerAspArgValIleAlaSerGlyLysPheIleGluGlyPheGlnGln 152
Qy 379 -----GCACCTGACTGTGCAATAACCGTA 402
Db 153 AlaLysLeuAlaAspProGlyAlaThrAsnArgAlaAlaProAlaIleSerValIle 172
Qy 403 CATACCAGGCAGATACGTCCAGTCCGATCCGCTTTTAAATCCTCTTAAACCTGGCGTT 462
Db 173 ---ProGluSerGluThrPheAsn-----AsnThrLeuAspHisGlyVal 186
Qy 463 TGCCAACTGGATAACGCGAACGTGACTGACGCGATCCTCAGCAGGCGAGGAGGTCAATT 522
Db 187 Cys-----ThrLysPheGluAlaSerGlnLeuGlyAspGluVal 199
Qy 523 ---GCTGACTTTACCGGGCATCGGCAACCGCGTTTCGGGAACCTGGACCGGTCCTTAAT 579
Db 200 AlaAlaAsnPheThr-----AlaLeuPhe 207
```



```
QY 580 TTTCGGCAATCAACTGTGCTTAAGCTGTAAGACGAGGAGGAAAGCTGCTTCAATACG 639
Db 208 AlaPro-----AspIleArgAlaArgAlaGluLysHisLeuProGlyValThrLeuThr 225
QY 640 CAGGCATTACCATCGGAATC-----AAGGTGAGCGCCGACCAATGCTCATTAACCGGT 693
Db 226 AspGluAspValValSerLeuMetAspMetCysSerPheAspThrValAlaAargThrSer 245
QY 694 GCGGTAAAGCTCGCATCAATGTCAGCGGATATTTCTCTGCAACAGACACAGGGAATG 753
Db 246 AspAlaSerGlnLeuSerProPheCysGlnLeuPheThrHisAsnGlu----- 261
QY 754 CCGGAGCCGGGTGGGAAGATCACCAGTTCACACAGTGGACACCTTGCTTAAGTTTG 813
Db 262 -----TTP-----LysLysTyrAsnTyrLeuGlnSerLeu 271
QY 814 CATACCGCGCAATTTATTTGTTACACGCGGAGGTGCGCGACGCGCGCAC 873
Db 272 -----GlyLysTyrTyrGlyTyrGlyAlaGlyAsnProLeuGlyProAlaGlnGlyLe 289
QY 874 CCGTTATTTGATTTGATCGGACGCTTGACGCCCTCCACCGCAAAACAG----- 927
Db 290 GlyPheThrAsnGluLeuIleAlaArgLeuThrArgSerProValGlnAspHisThrSer 309
QY 928 -----GCGTATGTTGTGACATTACCCACTTCAGTACTGTTTATTGCG 969
Db 310 ThrAsnSerThrLeuValSerAsnProAlaThrPheProLeuAsnAlaThrMetTyrVal 329
QY 970 -----GGACAGCATATACTCGCAATCTCGCGCGGCACTGGAGCTCAACTGCA 1023
Db 330 AspPheSerHisAspAsnSerMetValSerIlePheAlaLeu----- 344
QY 1024 CTTCGCCGTGAGCGGATAACACGCCCGGAGTGGTGAACCTGGTGTGAACTGGCGGT 1083
Db 345 -----GlyLeuTyrAsnGlyThrGluProLeuSerArgThrSerValGluSerAlaLys 362
QY 1084 CCGCTAAGCAT---AACAGCCAGTGGAT-----CAGGTTTCGCTGCTTCCAG 1131
Db 363 GluLeuAspGlyTyrSerAlaSerTrpValValProPheGlyAlaAargAlaTyrPheGlu 382
QY 1132 ACTTACAGCAGATGCTGATAAACCGCGCTGCTATTAATACGCCCGCGGAGAGGTG 1191
Db 383 ThrMetGlnCysLysSerGluLysGluProLeu---ValArgAlaLeuIleAsnAspArg 401
QY 1192 AAACCTGACCTGGCAGCATGTGAGAGCGAAATCGCGCAGGCATGTTGTTGCGCAGGT 1251
Db 402 ValValProLeuHisGlyCysAspVal---AspLysLeuGlyArgCysLysLeuAsnAsp 420
QY 1252 TTACGCAATCGTGAATGAAGCAGC 1278
Db 421 PheValLysGlyLeuSerTrpAlaAarg 429
```

RESULT 10

```
US-10-062-848-3
; Sequence 3, Application US/10062848
; Patent No. 6734004
; GENERAL INFORMATION:
; APPLICANT: KOSTREWA, Dirk
; APPLICANT: PASAMONTES, Luis
; APPLICANT: TOMSCHI, Andrea
; APPLICANT: van LOON, Adolphus
; APPLICANT: VOGEL, Kurt
; APPLICANT: WYSS, Markus
; TITLE OF INVENTION: MODIFIED PHYTASES
; FILE REFERENCE: Modified Phytases
; CURRENT APPLICATION NUMBER: US/10/062,848
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 09/044,718
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: EP 97810175.6
; PRIOR FILING DATE: 1997-03-25
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 3
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-062-848-3

Alignment Scores:
Pred. No.: 0.00207 Length: 439
Score: 119.00 Matches: 102
Percent Similarity: 36.67% Conservative: 70
Best Local Similarity: 21.75% Mismatches: 183
Query Match: 5.07% Indels: 114
DB: 4 Gaps: 24
```

US-10-034-985-1 (1-1323) x US-10-062-848-3 (1-439)

```
QY 19 CCATTTTATCTCTCTGATTCGTTAAACCCCGCAATCTGCAATTCGCTCAGATGAGCCG 78
Db 26 ProPhePheSerLeuGluAspGluLeuSerValSerLysLeuProLys----- 42
QY 79 GAGCTGAAGCTGGAAAGTGTGATTTGTCAGTCGTGTCATGTCGTGTCGTGTCCTCAACC-- 135
Db 43 AspCysArgIleThrLeuValGlnValLeuSerArgHisGlyAlaArgTyrProThrSer 62
QY 136 -----AAGCCACGCAACTGATGTCAGGATGTCACCCAGCAGCA----- 174
Db 63 SerLysSerLysLysTyrLysLysLeuValThrAlaIleGlnAlaAsnAlaThrAspPhe 82
QY 175 -----TGGCCAACTGCGCGGTAATAAAGCTGGTGG-----CTGACA 210
Db 83 LysGlyLysPheAlaPheLeuLysThrTyrAsnTyrThrLeuGlyAlaAspAspLeuThr 102
QY 211 CCCGNGGTGTCGATGATTAATCGCTATCTCGGA-----CATTACCAACGCCAGCGCTG 264
Db 103 ProPheGlyGlnGlnLeuValAsnSerGlyLysPheTyrGlnArgTyrLysAla 122
QY 265 GTAGCCGACGATGTCGTGGGAAAAAGGCGTCGCCGAGTCTGTCAGTCGCGGATATT 324
Db 123 LeuAlaArgSerValPro-----PheIleArg 132
QY 325 GCTGATGTCGACGAGGTACCCGTAACAGCGGAAGCCTTCCGCCGCGGCTG----- 378
Db 133 AlaSerGlySerAspArgValIleAlaSerGlyGluLysPheIleGluGlyPheGln 152
QY 379 -----GCACCTGACTGTGCAATAACCGTA 402
Db 153 AlaLysLeuAlaAspProGlyAlaThrAsnArgAlaAlaProAlaIleSerValIle 172
QY 403 CATACCCAGCAGATAGTCCTCCGATCCGATCCGTTATTTATCTCTAAAAACCTGGCGTT 462
Db 173 ---ProGluSerGluThrPheAsn-----AsnThrLeuAspHisGlyVal 186
QY 463 TGCCAACTGGATAACCGAAGCTGACTGACGCGATCCTCAGCAGGCGGAGGTCAT 522
Db 187 Cys-----ThrLysPheGluAlaSerGlnLeuGlyAspGluVal 199
QY 523 ---GCTGACTTTTACCGGCGATCGGCAACCGCGTTTCGCGAAGCTGGAACGGGTCTTAAT 579
Db 200 AlaAlaAsnPheThr-----AlaLeuPhe 207
QY 580 TTTCGCAATCAAACTTGTGCTTAAACGTGAGAAACAGACGACCAAGCTGCTTCAATACG 639
Db 208 AlaPro-----AspIleArgAlaArgAlaGluLysHisLeuProGlyValThrLeuThr 225
QY 640 CAGGCATTACCATCGGAATC-----AAGGTGAGCGCCGACCAATGCTCATTAACCGGT 693
Db 226 AspGluAspValValSerLeuMetAspMetCysSerPheAspThrValAlaAargThrSer 245
QY 694 GCGGTAAAGCTCGCATCAATGTCAGCGGATATTTCTCTGCAACAGACACAGGGAATG 753
Db 246 AspAlaSerGlnLeuSerProPheCysGlnLeuPheThrHisAsnGlu----- 261
QY 754 CCGGAGCCGGGTGGGAAGGATCACCGATTACACCGATTCACCGATGGGAACACCTTGAAGTTTG 813
```

```
Db      262  ---LysLysTyrAsnTyrLeuGlnSerLeu 271
Qy      814  CATAACCGCGCAATTTATTTGCTACAAACGACGAGAGTTGCCCGCAGCGCGCCACC 873
Db      272  -----GlyLysTyrGlyTyrGlyAlaGlyAsnProLeuGlyProAlaGlnGlyIle 289
Qy      874  CCGTTATTGATTGATCATGGAGCGCTTCAGCGCCCATCCAGCGCAAAACAG----- 927
Db      290  GlyPheThrAsnGluLeuIleAlaArgLeuThrArgSerProValGlnAspHisThrSer 309
Qy      928  -----GGGTATGTGTGACATTTACCCACTTCAGTACTGTTTATTGGC 969
Db      310  ThrAsnSerThrLeuValSerAsnProAlaThrPheProLeuAsnAlaThrMetTyrVal 329
Qy      970  -----GGACAGCATTAATCTGGCAATCTCGCGCGCAGCTGGAGCTCAACTGGACG 1023
Db      330  AspPheSerHisAspAsnSerMetValSerIlePheAlaLeu----- 344
Qy      1024  CTTCCCGGTGAGCGGATACAGCGCGCAGGTGGTGAACCTGGTGTGAAACGTCGGCGT 1083
Db      345  -----GlyLeuTyrAsnGlyThrGluProLeuSerArgThrSerValGluSerAlaLys 362
Qy      1084  CGCTTAAGCGAT---AACACCGAGTGGATT-----CAGGTTTCGCTGGTCTTCCAG 1131
Db      363  GluLeuAspGlyTyrSerAlaSerTrpValValProPheGlyAlaArgAlaTyrPheGlu 382
Qy      1132  ACTTTACAGCAGATGCGTGATATAACCGCCGCTGCTATTAATACGCGCGCGGAGAGGTG 1191
Db      383  ThrMetGlnCysLysSerGluLysGluProLeu---ValArgAlaLeuIleAsnAspArg 401
Qy      1192  AAAGTACCTGCGCAGATGTGAAGACGAAATGCGCAGGCGCATGCTGTTGGTGGCAGGT 1251
Db      402  ValValProLeuHisGlyCysAspVal---AspLysLeuGlyArgCysLysLeuAsnAsp 420
Qy      1252  TTTACGCAATCGTGAATGAAGCACGC 1278
Db      421  PheValLysGlyLeuSerTrpAlaArg 429

RESULT 11
US-09-684-855-106
; Sequence 106, Application US/09684855
; Patent NO. 6599735
; GENERAL INFORMATION:
; APPLICANT: P. Hoffmann-La Roche AG
; TITLE OF INVENTION: CONTINUOUS FERMENTATION PROCESS
; FILE REFERENCE: C38435/111692
; CURRENT APPLICATION NUMBER: US/09/684,855
; CURRENT FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: EP 00121663.9
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: EP 99120289.6
; PRIOR FILING DATE: 1999-10-11
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 106
; LENGTH: 440
; TYPE: PRT
; ORGANISM: A. fumigatus 13073
US-09-684-855-106

Alignment Scores:
Pred. No.:      0.00207      Length:      440
Score:          119.00      Matches:     102
Percent Similarity: 36.67%      Conservative: 70
Best Local Similarity: 21.75%      Mismatches:  183
Query Match:      5.07%      Indels:     114
DB:              4          Gaps:        24

US-10-034-985-1 (1-1323) x US-09-684-855-106 (1-440)
Qy      i9  CCATTTTATCTCTTCTGATTCGTTAAACCGCGCAATCTCGATTTCGCTCAGAGTGAGCCG 78
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
          |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

```
Db      27  ProPhePheSerLeuGluAspGluLeuSerValSerSerLysLeuProLys----- 43
Qy      79  GACCTGAAGCTGAAAGTGTGGTGAATTGTCAGTCGTCATGGTGTGGTGTGCTCAACG--- 135
Db      44  AspCysArgIleThrLeuValGlnValLeuSerArgHisGlyAlaArgTyrProThrSer 63
Qy      136  -----AAGCGCCACGCAACTGATGTGAGGATGTCAACCCACGACGCA----- 174
Db      64  SerLysSerLysLysTyrLysLysLeuValThrAlaIleGlnAlaAsnAlaThrAspPhe 83
Qy      175  -----TGGCAACCTGGCCCGGTAAACCTGGGTGG-----CTGACA 210
Db      84  LysGlyLysPheAlaPheLeuLysThrTyrAsnTyrThrLeuGlyAlaAspAspLeuThr 103
Qy      211  CCGCGNGGTGGTGCAGCTAATCGCTATCTCGA-----CATTACCAACGCGCGCTG 264
Db      104  ProPheGlyGlnGlnGlnLeuValAsnSerGlyIleLysPheTyrGlnArgTyrLysAla 123
Qy      265  GTAGCCGAGCGGATTTGCTGCGCAAAAGGGTGTCCCGCAGTCTGTCAGTCTCGGATTATT 324
Db      124  LeuAlaArgSerValValPro-----PheIleArg 133
Qy      325  GCTGATGTGCAGAGGTATCCCGTAAACAGGCGAAGCCTTCCCGCGCGGCTG----- 378
Db      134  AlaserGlySerAspArgValIleAlaserGlyGluLysPheIleGluGlyPheGlnGln 153
Qy      379  -----GCACCTGACTGTGCAATAACCGTA 402
Db      154  AlaLysLeuAlaAspProGlyAlaThrAsnArgAlaAlaProAlaIleSerValIleIle 173
Qy      403  CATACCAGGCGAGATACGTCACGTCCTCCGATCCGTTATTTAACTCTCTAAACCTGGCGTT 462
Db      174  ---ProGluSerGluThrPheAsn-----AsnThrLeuAspHisGlyVal 187
Qy      463  TGCACAACTGGATAACCGAACGACGTGACGCGATCTCAGAGGCGAGGAGGTCAATT 522
Db      188  Cys-----ThrLysPheGluAlaserGlnLeuGlyAspGluVal 200
Qy      523  ---GCTGACTTTTACCGGCGATCGGCAACGCGGCTTTCGGGAACCTGGAAACGGGTCTAAT 579
Db      201  AlaAlaAsnPheThr-----AlaLeuPhe 208
Qy      580  TTTCCCGCAATCAAACTTGTGCTTAAACGTCGAGAAACAGGACGAAAGCTGTTTCATTAACG 639
Db      209  AlaPro-----AspIleArgAlaArgAlaGluLysHisLeuProGlyValThrLeuThr 226
Qy      640  CAGCGAATTCACCTCGGAACCTC-----AAGTGAGCGCGCGACAAATGCTCTCATTAACCGGT 693
Db      227  AspGluAspValValSerLeuMetAspMetCysSerPheAspThrValAlaArgThrSer 246
Qy      694  GCGGTAAGCTCGCATCAATGTCGCGGAGATATTTCTCTGCAACAGCACACGAGGAATG 753
Db      247  AspAlaserGlnLeuSerProPheCysGlnLeuPheThrHisAsnGlu----- 262
Qy      754  CCGGAGCGGGGTGGGAGGAGTACCGGATTCACACAGTGGGAACACCTTGTCTAAGTTTG 813
Db      263  -----Trp-----LysLysTyrAsnTyrLeuGlnSerLeu 272
Qy      814  CATAACGCGCAATTTTATTGTTGTACAAACGCGCAGAGAGTTGCCCGCAGCGCGCCACC 873
Db      273  -----GlyLysTyrTyrGlyTyrGlyAlaGlyAsnProLeuGlyProAlaGlnGlyIle 290
Qy      874  CCGTTATTGATTGATCATGGCAGCGTTTCACCGCCCATCCCGCGCAAAACAG----- 927
Db      291  GlyPheThrAsnGluLeuIleAlaArgLeuThrArgSerProValGlnAspHisThrSer 310
Qy      928  -----CGGTATGTGTGACATTAACCCACTTCAGTACTGTTTATTGGC 969
Db      311  ThrAsnSerThrLeuValSerAsnProAlaThrPheProLeuAsnAlaThrMetTyrVal 330
Qy      970  -----GGACAGCATTAATCTGGCAATCTCGGCGCGCGCAGCTGGAGCTCAACTGACG 1023
Db      331  AspPheSerHisAspAsnSerMetValSerIlePhePheAlaLeu----- 345
```


Qy	1024	CTTCCCGGTGAGCCGGATAACACC	CGCCGCGCAGGTGGTAACGCTGC	GCTGTCGTTGAACGCTGC	GCT	1083
Db	346	-----GlyLeuTyAenGlyThrGlup	ProLeuSerArgThrServalGluSerAla	Lys		363
Qy	1084	CGGCTAACCGAT---AAACGCCAGT	GATT-----CAGGTTCCTGGTGCT	TTC	CAG	1131
Db	364	GluleuAspGlyTySerAlaSerTrp	VaiValProPheGlyAlaAgaAlaTy	Phe	Glu	383
Qy	1132	ACTTTACAGCAGATGCGTGTATAA	ACCCGCTGTCATTAAATACGCCCG	CAGAGAGTG		1191
Db	384	ThrMetGlnCynLysSerGluLysGlu	ProLeu---ValargAlaLeulleAsn	Asp	Arg	402
Qy	1192	AACTGACCTCGCAGCATGGAAGAC	GCGAAATCCGCGAGGCATGTGTT	CGTTGGCAGGT		1251
Db	403	ValValProLeuHisGlyCysAspVal	---AspLysLeuGlyArgCysLys	Leu	Asn	421
Qy	1252	TTTTACGCAATCGTAATGAAGACGC				1278
Db	422	PheValLysGlyLeuSerTrdAlaArg				430

RESULT 12

US-09-684-855-128
; Sequence 128, Application US/09684855
; Patent No. 6599735
; GENERAL INFORMATION:
; APPLICANT: F. Hoffmann-La Roche AG
; TITLE OF INVENTION: CONTINUOUS FERMENTATION PROCESS
; FILE REFERENCE: C38435/111692
; CURRENT APPLICATION NUMBER: US/09/684,855
; CURRENT FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: EP 00121663.9
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: EP 99120289.6
; PRIOR FILING DATE: 1999-10-11
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 128
; LENGTH: 440
; TYPE: PRT
; ORGANISM: A. fumigatus 13073
US-09-684-855-128

Alignment Scores:	
Pred. No.:	0.00207
Score:	119.00
Percent Similarity:	36.67%
Best Local Similarity:	21.75%
Query Match:	5.07%
DB:	4
Length:	440
Matches:	102
Conservative:	70
Mismatches:	183
Indels:	114
Gaps:	24

US-10-034-985-1 (1-1323) x US-09-684-855-128 (1-440)

19	Qy	CCATTTTATCTCTTCTGATTCCGTTAAACCCGCAATCTGCATTCGCTCAGAGTCGAGCGG	78
27	Db	ProPheSerLeuGluAspGluLeuSerValSerLysLeuProLys	43
79	Qy	GAGCTGAAGCTGGAAAGTGTGTGATTGTCAGTCGTCATGGTGGTGGTCTCCAACC	135
44	Db	AspCysArgIleThrLeuValGlnValLeuSerArgHisGlyAlaArgTyrProThrSer	63
136	Qy	-----AAGGCCACCGCACTGATGTCAGGATGTCACCCAGACGCA-----	174
64	Db	SerLysSerLysTyrsTyrsLysLeuValThrAlaIleGlnAlaAsnAlaThrAspPhe	83
175	Qy	-----TGGCCAACCTGGCCGGTAAACCTGGTGG-----CTGACA	210
84	Db	LysGlyLysPheAlaPheLeuLysThrTyrAsnTyrThrLeuGlyAlaAspAspLeuThr	103
211	Qy	CCGCGNGGTGTGTGACTAATCGCCTATCTCGGA-----CATTAACAACGCCAGCGTCTG	264
104	Db	ProPheGluGluGlnLeuValAsnSerGlyIleLysPheTyrGlnArgTyrLysAla	123

```
Db 422 PheVallysGlyLeuSerTrpAlaArg 430
|||||
US-09-684-855-151
; Sequence 151, Application US/09684855
; Patent No. 6599735
; GENERAL INFORMATION:
; APPLICANT: F. Hoffmann-La Roche AG
; TITLE OF INVENTION: CONTINUOUS FERMENTATION PROCESS
; FILE REFERENCE: C38435/111692
; CURRENT APPLICATION NUMBER: US/09/684,855
; CURRENT FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: EP 00121663.9
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: EP 99120289.6
; PRIOR FILING DATE: 1999-10-11
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 151
; LENGTH: 440
; TYPE: PRT
; ORGANISM: A. fumigatus 32722
US-09-684-855-151

Alignment Scores:
Pred. No.: 0.00207 Length: 440
Score: 119.00 Matches: 102
Percent Similarity: 36.67% Conservative: 70
Best Local Similarity: 21.75% Mismatches: 183
Query Match: 5.07% Indels: 114
DB: Gaps: 24

US-10-034-985-1 (1-1323) x US-09-684-855-151 (1-440)
QY 19 CCATTTTATCTCTTCGATTCGCTTAACCCCGCAATCTCATTCGCTCAGAGTGAGCCG 78
|||||
Db 27 ProPheSerLeuGluAspGluLeuSerValSerSerLysLeuProLys----- 43
QY 79 GAGCTGAAGCTGGAAGCTGGTGGTATGTCAGTCGTCATGGTGGTGGTCCACCC- 135
|||
Db 44 AspCysArgileThrLeuValGlnValLeuSerArgHisGlyAlaArgTyrPr-oThrSer 63
|||
QY 136 -----AAGCCACGCACTGATGAGGATGTCAACCCAGACGCA----- 174
|||
Db 64 SerLysSerLysTyrLysLeuValThrAlaIleGlnAlaAsnAlaThrAspPhe 83
|||
QY 175 -----TGGCCAACTGGCCCGGTAAACTGGTGG-----CTGACA 210
|||
Db 84 LysGlyLysPheAlaPheLeuLysThrTyrAsnTyrThrLeuGlyAlaAspLeuThr 103
|||
QY 211 CCGCGNGGTGGTGAAGTAATCGCTATCTCGGA-----CATTAACCAACCGCGCTCTG 264
|||
Db 104 ProPheGlyGluGlnGlnLeuValAsnSerGlyIleLysPheTyrGlnArgTyrLysAla 123
|||
QY 265 GTAGCCGACGATGTCGCGCAAAAGGGTGGCCGAGTCTGTCAGGTGCGGATATT 324
|||
Db 124 LeuAlaArgSerValValPro-----PheIleArg 133
|||
QY 325 GCTGATGTCAGCAGGCTACCGGTAAACAGCGCAAGCCTTCGCGCGCGGCTG----- 378
|||
Db 134 AlaSerGlySerAspArgValIleAlaSerGlyGluLysPheIleGluGlyPheGlnGln 153
|||
QY 379 -----GCACCTGACTGTGCAATAACCGTA 402
|||
Db 154 AlaLysLeuAlaAspProGlyAlaThrAsnArgAlaAlaProAlaIleSerValIle 173
|||
QY 403 CATACCAGGACAGATACGTCAGTCCGATCCGTTATTTAATCTCTATAAACTGGCGTT 462
|||
Db 174 ---ProGluSerGluThrPheAsn-----AsnThrLeuAspHisGlyVal 187
|||
QY 463 TGCCAACTGATACCGCAACGTGATCTGACGCGATCTCAGAGGCGGAGGAGGTCAATT 522
|||
```

```
Db 188 Cys-----ThrLysPheGluAlaSerGlnLeuGlyAspGluVal 200
QY 523 ---GCTGACTTTTACCGGGCATCGCAAAACGGCGTTTCGCGAACTGGAACGGGTGCTTAAT 579
|||||
Db 201 AlaAlaAsnPheThr-----AlaLeuPhe 208
QY 580 TTTCGCAATCAAACTTTGTGCTTAAACGTCGAGAAACAGGACGAAAGCTGCTTCAATACG 639
|||
Db 209 AlaPro-----AspIleArgAlaArgAlaGluLysHisLeuProGlyValThrLeuThr 226
QY 640 CAGGCATTACCATCGGAATC-----AAGTGAGCGCCGACAAATGCTCTCAATCAACCGGT 693
|||
Db 227 AspGluAspValValSerLeuMetAspMetCysSerPheAspThrValAlaIleArgThrSer 246
QY 694 GCGGTAAGCCTCGCATCAATGTCGCGGAGATATTTCTCTCTGCAACAAAGCACACGGAATG 753
|||
Db 247 AspAlaSerGlnLeuSerProPheCysGlnLeuPheThrHisAsnGlu----- 262
QY 754 CCGGAGCGGGGTGGGGAAGGATCACCGATTTCACACAGTGGAAACACCTTTGCTAAGTTTG 813
|||
Db 263 -----Trp-----LysLysTyrAsnTyrLeuGlnSerLeu 272
QY 814 CATAACGCGCAATTTTATTGCTTACAAACGACCGCAGAGGTTGCCCGCAGCCGCGCAC 873
|||
Db 273 -----GlyLysTyrTyrGlyTyrGlyAlaGlyAsnProLeuGlyProAlaGlnGlyIle 290
QY 874 CCGTTATTGGATTGATCATGCGCAGCGTTGACGCCCATCCACCGCAAAACACAG----- 927
Db 291 GlyPheThrAsnGluLeuIleAlaArgLeuThrArgSerProValGlnAspHisThrSer 310
QY 928 -----GCGTATGGTGTGACATTAACCCACTTTCAGTACTGTTTATTATGCC 969
|||
Db 311 ThrAsnSerThrLeuValSerAsnProAlaThrPheProLeuAsnAlaThrMetTyrVal 330
QY 970 -----GGACACGATTAATCTCGCAAAATCTCGCGCGGCACCTGAGGACTCAACTGACG 1023
Db 331 AspPheSerHisAspAsnSerMetValSerIlePhePheAlaLeu----- 345
QY 1024 CTTCCCGGTGACCGGATAACACGCGCCGAGGTGGTGAACCTGTTTGAACGCTGGCGT 1083
|||
Db 346 -----GlyLeuTyrAsnGlyThrGluProLeuSerArgThrSerValGluSerAlaLys 363
QY 1084 CGCTTAAGCGAT---AACAGCCAGTGGATT-----CAGTTTCGCTGCTTCCAG 1131
|||
Db 364 GluLeuAspGlyTyrSerAlaSerTrpValValProPheGlyAlaArgAlaTyrPheGlu 383
QY 1132 ACTTTACAGCAGATGCGTGATAAACGCGCTGCTCATTAATAATACGCGCCGAGAGGTG 1191
|||
Db 384 ThrMetGlnCysLysSerGluLysGluProLeu---ValArgAlaLeuIleAsnAspArg 402
QY 1192 AAATCAGCCTGCGCAGATGTGAAGAGCGAAATGCGCAGGCGATGTTGTTGGCAGGT 1251
|||
Db 403 valValProLeuHisGlyCysAspVal---AspLysLeuGlyArgCysLysLeuAsnAsp 421
QY 1252 TTTCACGAAATCGTGAATGAAGCACGC 1278
|||
Db 422 PheVallysGlyLeuSerTrpAlaArg 430

RESULT 14
US-09-488-265B-6
; Sequence 6, Application US/09488265B
; Patent No. 6720174
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Martin
; APPLICANT: Lassen, Soren F
; TITLE OF INVENTION: Improved Phytases
; FILE REFERENCE: 5808.200-US
; CURRENT APPLICATION NUMBER: US/09/488,265B
; CURRENT FILING DATE: 2000-01-20
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 440
```

```

; TYPE: PRT
; ORGANISM: Aspergillus fumigatus 13073
US-09-488-265B-6

Alignment Scores:
Pred. No.: 0.00207 Length: 440
Score: 119.00 Matches: 102
Percent Similarity: 36.67% Conservative: 70
Best Local Similarity: 21.75% Mismatches: 183
Query Match: 5.07% Indels: 114
DB: 4 Gaps: 24

US-10-034-985-1 (1-1323) x US-09-488-265B-6 (1-440)
Qy 19 CCATTTTATCTCTTCTGATTCCGTTAAACCCCGCAATCTGCATTCGCTCAGAGTGAGCGG 78
Db 27 ProphePheSerLeuGluAspGluLeuSerValSerSerLeuProLys----- 43
Qy 79 GAGCTGAAGCTGGAAGTGTGGTGAATTCAGTCAGTCATGCTGTGCTGCTGCTCAACC--- 135
Db 44 AspCysArgIleThrLeuValGlnValLeuSerArgHisGlyAlaAArgTyrProThrSer 63
Qy 136 -----AAGGCCACCACTGATGTCAGGATGTCACCCAGACGCA----- 174
Db 64 SerLysSerLysTyrLysLeuValThrAlaIleGlnAlaAsnAlaThrAspPhe 83
Qy 175 -----TGCCCAACCTGGCGCGGTAAACTGGGTTGG-----CTGACA 210
Db 84 LysGlyLysPheAlaPheLeuLysThrTyrAsnTyrThrLeuGlyAlaAspAspLeuThr 103
Qy 211 CCGGNGGTGTGAGCTAATCGCTATCTCGGA-----CATTACCAACGCCAGCGTGTG 264
Db 104 ProPheGlyGluGlnGlnLeuValAsnSerGlyIleLysPheTyrGlnAArgTyrLysAla 123
Qy 265 GTACCCGACGATGCTGGCGGAAAGGGTCCCGCAGTCGTGTCAGGTCGCGATTATT 324
Db 124 LeuAlaArgSerValPro-----PheIleArg 133
Qy 325 GCTGATGTCGACGAGCTACCGTAAACAGCGGAAGCTTCGCGCGGCGCTG----- 378
Db 134 AlaSerGlySerAspArgValIleAlaSerGlyLysPheIleGlyLysPheGlnGln 153
Qy 379 -----GCACCTGACTGTGCAATTAACCGTA 402
Db 154 AlaLysLeuAlaAspProGlyAlaThrAsnArgAlaAlaProAlaIleSerValIle 173
Qy 403 CATACCCAGGAGATAGTCCAGTCCCGATCCGTTATTATCTCTAAACCTGCGGCTT 462
Db 174 ---ProGluSerGluThrPheAsn-----AsnThrLeuAspHisGlyVal 187
Qy 463 TGCCAACTGGATAACGGGAACGTGACTGACGCGATCCTCAGCAGCGGAGGAGGTCATTT 522
Db 188 Cys-----ThrLysPheGluAlaSerGlnLeuGlyAspGluVal 200
Qy 523 ---GCTGACTTTACGGGATCGGCAACAGCGGTTTCGCGAAGTGGAAACGGGTCCTTAAT 579
Db 201 AlaAlaAsnPheThr-----AlaLeuPhe 208
Qy 580 TTTCGGCAATCAACTTGTGCTTAAAGCTGAGAAACAGGACGAAAGCTGTTCATTAAACG 639
Db 209 AlaPro-----AspIleArgAlaAlaGluLysHisLeuProGlyValThrLeuThr 226
Qy 640 CAGGCATTACCATCGGAATCTC-----AAGGTGAGCGCCGACAAATGTCATTAAACCGGT 693
Db 227 AspGluAspValValSerLeuMetAspMetCysSerPheAspThrValAlaAArgThrSer 246
Qy 694 GCGGTAAAGCTCGCATCAATGCTGACGAGATATTTCTCTGCAACCAAGCACAGGGAATG 753
Db 247 AspAlaSerGlnLeuSerProPheCysGlnLeuPheThrHisAsnGlu----- 262
Qy 754 CCGGAGCGGGGTGGGAGGATCACCAGATTACACAGGTGGGAACACCTGCTGAAGTTTG 813
Db 263 -----Trp-----LysLysTyrAsnTyrLeuGlnSerLeu 272

RESULT 15
US-09-044-718-12
; Sequence 12, Application US/09044718
; Patent No. 6391605
; GENERAL INFORMATION:
; APPLICANT: KOSTREWA, Dirk
; APPLICANT: PASAMONTES, Luis
; APPLICANT: TOMSCHY, Andrea
; APPLICANT: van LOON, Adolphus
; APPLICANT: VOGEL, Kurt
; APPLICANT: WYSS, Markus
; TITLE OF INVENTION: MODIFIED PHYTASES
; FILE REFERENCE: Modified Phytases
; CURRENT APPLICATION NUMBER: US/09/044,718
; CURRENT FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: EP 97810175.6
; PRIOR FILING DATE: 1997-03-25
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-09-044-718-12

Alignment Scores:
Pred. No.: 0.00209 Length: 449
Score: 119.00 Matches: 102
Percent Similarity: 36.67% Conservative: 70
Best Local Similarity: 21.75% Mismatches: 183
Query Match: 5.07% Indels: 114
DB: 3 Gaps: 24

US-10-034-985-1 (1-1323) x US-09-044-718-12 (1-449)
Qy 19 CCATTTTATCTCTTCTGATTCCGTTAAACCCCGCAATCTGCATTCGCTCAGAGTGAGCGG 78
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 11, 2005, 23:08:45 ; Search time 208.5 Seconds

(without alignments)
4908.240 Million cell updates/sec

Title: US-10-034-985-1

Perfect score: 2349

Sequence: 1 atgaagcgtatcttaatcc.....atcaccatcaccatcaactaa 1323

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 4211384

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DBV=xlp
-O=/cgn2.1/USPTO.spool.p/US10034985/runat.10032005.164239.11943/app.query.fasta_1.1479
-DB=A_Geneseq_16Dec04 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptp -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10034985 @CGN 1.1 224 @runat.10032005.164239.11943 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DVT TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_16Dec04:.*
1: Geneseq1980s:.*
2: Geneseq1990s:.*
3: Geneseq2000s:.*
4: Geneseq2001s:.*
5: Geneseq2002s:.*
6: Geneseq2003as:.*
7: Geneseq2003bs:.*
8: Geneseq2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2302	98.0	440	4 AAB37892	Aab37892 Escherich
2	2302	98.0	440	5 AAE22836	Aae22836 Escherich
3	2302	98.0	440	5 AAE15806	Aae15806 Escherich
4	2302	98.0	440	6 ADA19446	Ada19446 E. coli B
5	2302	98.0	440	8 ADO50292	Ado50292 Escherich
6	2282	97.1	440	8 ADO50304	Ado50304 Escherich
7	2279	97.0	440	8 ADO50302	Ado50302 Kangaroo
8	2235	95.1	432	3 AAB36257	Aab36257 Lama2/APP
9	2235	95.1	432	3 AAB36261	Aab36261 R15/APPA
10	2235	95.1	432	3 AAB36262	Aab36262 SV40/APPA

11	2235	95.1	432	3 AAB36259	Aab36259 R15/APPA
12	2235	95.1	432	3 AAB36258	Aab36258 R15/APPA
13	2235	95.1	432	3 AAB36263	Aab36263 Lama2/APP
14	2235	95.1	432	3 AAB36260	Aab36260 R15/APPA
15	2235	95.1	432	3 AAU77775	Aau77775 Phytase a
16	2235	95.1	432	5 AAE15807	Aae15807 Escherich
17	2235	95.1	432	5 ABP53625	Abp53625 Acid phos
18	2235	95.1	432	6 ADA19450	Ada19450 E. coli K
19	2235	95.1	432	6 ABU15427	Abu15427 Protein e
20	2235	95.1	432	7 ADC87745	Adc87745 Escherich
21	2235	95.1	432	8 ADL16129	Adl16129 Escherich
22	2235	95.1	432	8 ADO50298	Ado50298 Escherich
23	2230	94.9	432	4 AAE02631	Aae02631 E. coli a
24	2230	94.9	432	8 ADL16131	Adl16131 Shigella
25	2224	94.7	432	4 AAE02634	Aae02634 E. coli a
26	2224	94.7	432	8 ADL16130	Adl16130 Shigella
27	2223	94.6	432	4 AAE02635	Aae02635 E. coli a
28	2221	94.6	432	3 AAY94753	Aay94753 E. coli a
29	2221	94.6	432	6 ABR42161	Abr42161 Escherich
30	2210	94.1	432	4 AAE02632	Aae02632 E. coli a
31	2198	93.6	432	7 ADL02198	Adl02198 Appa. 5/2
32	2197	93.5	432	6 ABR42162	Abr42162 Escherich
33	2194.5	93.4	423	2 AAY01513	Aay01513 An Escher
34	2188	93.1	436	7 ADC87743	Adc87743 Modified
35	2171	92.4	432	8 ADO50300	Ado50300 Escherich
36	2141	91.1	430	5 AAE15808	Aae15808 Escherich
37	2141	91.1	430	6 ADA19451	Ada19451 E. coli K
38	2137.5	91.0	440	8 ADL16120	Adl16120 E. coli A
39	2131.5	90.7	440	8 ADL16125	Adl16125 Modified
40	2129.5	90.7	440	8 ADL16126	Adl16126 Modified
41	2129.5	90.7	440	8 ADL16122	Adl16122 Modified
42	2125.5	90.5	440	8 ADL16124	Adl16124 Modified
43	2121.5	90.3	440	8 ADL16128	Adl16128 Modified
44	2120.5	90.3	440	8 ADL16127	Adl16127 Modified
45	2112.5	89.9	440	8 ADL16123	Adl16123 Modified

ALIGNMENTS

RESULT 1

AAB37892

ID AAB37892 standard; protein; 440 AA.

XX AAB37892;

AC AAB37892;

DT 07-MAR-2001 (first entry)

XX Escherichia coli B phytase enzyme.

DE Escherichia coli B phytase enzyme.

XX Escherichia coli B; phytase enzyme; anabolic; phytate digestion;

KW nutrition.

XX Escherichia coli.

OS Escherichia coli.

XX WO2000071728-A1.

XX 30-NOV-2000.

XX 25-MAY-2000; 2000WO-US014846.

XX 25-MAY-1999; 99US-00318528.

XX (DIVE-) DIVERSA CORP.

XX Short JM, Kretz KA;

XX WPI; 2001-112081/12.

XX N-PSDB; AAC88885.

XX Improving the nutritional value of phytate-containing foodstuffs, using

PT phytase enzymes which catalyze the liberation of inorganic phosphate from

PT the phytates.

XX

PS	Claim 1; Fig 1; 147pp; English.	
XX	The present sequence is a phytase enzyme from Escherichia coli B. The	
CC	enzyme catalyses the liberation of inorganic phosphate from the phytate	
CC	in phytate-containing foodstuffs and can thus be used to improve the	
CC	nutritional value of phytate rich ingredients	
XX	Sequence 440 AA;	
SQ		
	Alignment Scores:	
	Pred. No.: 8,77e-230 Length: 440	
	Score: 2302.00 Matches: 440	
	Percent Similarity: 100.00% Conservative: 0	
	Best Local Similarity: 100.00% Mismatches: 0	
	Query Match: 98.00% Indels: 0	
	DB: 4 Gaps: 0	
	US-10-034-985-1 (1-1323) x AAB37892 (1-440)	
QY	1 ATGAAAGCGATCTTAATCCCATTTTATCTCTCTGATTCCGTTAAACCCCGCAATCTGCA 60	261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db	1 MetLysAlaIleLeuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20	841 CGCAGCCCGAGAGTTGCGCGCAGCGCCGACCCCGTATTGATTTGATTCATGGAGCG 900
QY	61 TTGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGTGATGTGCTCATGTCTCATGT 120	281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuLeuMetAlaAla 300
Db	21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40	901 TTGACGCCCGCATCCACCGCAAAACAGGCGTATGTTGTGACATTATCCCATCTTCAGTACTG 960
QY	121 GTGCTGCTCCACACCGAGCGCAGCACTGATCAGAGTGTACCCAGCAGCGATCGCCA 180	301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db	41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60	961 TTTATTTCGGGACACGATTAATCTGGCAAAATCTCGGGCGGCACCTGGAGCTCAACTGG 1020
QY	181 ACCTGGCCGTAATACTGGCTTGGCTGACACCCCGNGGTGGTGGAGTAAATCGCTATCTC 240	321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db	61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80	1021 ACGCTTCCCGTCCAGCGGATACACCGCGCAGGTGGTGAACCTGGTGTGTTGAACGCTGG 1080
QY	241 GGACATTACCAACCGCAGCGCTGTGGTAGCGAGCGATGTGCGGAAAAAGGGCTGCCG 300	341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db	81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100	1081 CGTGGCTAAGCGATAACACGAGTTCAGTTTCGCTGCTTCTCCAGACTTTACAG 1140
QY	301 CAGCTGCTCAGTCCGATTATGCTGATGTCGAGCGGTACCGGTAAACAGCGCGA 360	361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db	101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120	1141 CAGATCGGTGATAAAGCGCGCTGCTCATTTAAATACCGCGCCGAGAGGTGAACCTGACC 1200
QY	361 GCCTTGGCCCGCGGTGGCACCTGACTGTGCAATAACCGTACATACCCAGCGAGATACG 420	381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db	121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140	1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCGATGTGTCGTGGCAGGTTTACGCAA 1260
QY	421 TCAGTCCCGATCCGTTATTATTCCTTAAACCTGGCGTTTGGCAACTGGATAACGCG 480	401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db	141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160	1261 ATCGTCAATGAAGCACGATACCGCGTGCAGTTTGAGATCTCATCCATCACCATCAC 1320
QY	481 AAGCTGACTGACCGATCCCTCAGCAGCGCAGGAGGTCAATTGCTGACTTTACCGGCAT 540	421 IleValAsnGluAlaAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis 440
Db	161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180	
QY	541 CGCAAAACGGCGTTTCGCAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTCG 600	RESULT 2
Db	181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200	AAEZ2836
QY	601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAAACGGTACCATTCATCGAACTC 660	ID AAE22836 standard; protein; 440 AA.
Db	201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220	XX AAE22836;
QY	661 AAGCTGAGCGCCACATGCTCATTAACCGGTGCGTAAAGCTTCCATCAATGCTGACG 720	XX 21-AUG-2002 (first entry)
Db	221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240	DE Escherichia coli phytase protein.
QY	721 GAGATATTTCTCTGCAACAGCACAGGGAATCGCGAGCGCGGTGGGGAAGGATCACCC 780	KW Dietary aid; biocompatible composition; therapeutic; digestive tract;
Db	241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260	KM foodstuff; digestion; phytase; enzyme.
QY	781 GATTACACCGTGGAAACACTTGTCTGAAGTTTGGATTAACCGCAATTTATTTGCTACAA 840	XX Escherichia coli.
		OS WO200189317-A2.
		XX 29-NOV-2001.
		XX 15-MAY-2001; 2001WO-US015764.
		XX 25-MAY-2000; 2000US-00580937.
		XX (DIVE-) DIVERSA CORP.
		XX Short JM, Kretz KA, O'donoghue E;
		XX WPI; 2002-164149/21.
		XX N-PSDB; AAD36473.
		XX New dietary aids comprising sustained release biocompatible compositions,
		PT comprise agent that assists in digestion, useful for delivering enzymes,
		PT therapeutics, medicine or agents to an organism.
		XX Claim 5; Fig 1; 89pp; English.
		XX The present invention relates to novel dietary aids comprising sustained
		CC release biocompatible composition which comprises an agent (enzymes such
		CC as phytase, amylase, esterase, protease) that assists in digestion. The

CC biocompatible composition is effective upon oral consumption and release
 CC in the digestive tract of a subject. The dietary aids are useful for
 CC delivering enzymes, therapeutics, medicine and agents to an organism. The
 CC use of enzymes and other agents in digestive aids of livestock or
 CC domesticated animals not only improves the animal's health and life
 CC expectancy but also assists in increasing the health of livestock or in
 CC the production of foodstuffs from livestock. The present sequence is
 CC Escherichia coli phytase protein
 XX

SQ Sequence 440 AA;

Alignment Scores:

Prod. No.: 8 77e-230 Length: 440
 Score: 2302.00 Matches: 440
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.00% Indels: 0
 DB: 5 Gaps: 0

US-10-034-985-1 (1-1323) x AAE22836 (1-440)

QY 1 ATGAAGCGATCTTAATCCATTTTATCTCTCTGATTCGTTAAACCCCGCAATCTGCA 60
 DB 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 QY 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAAGTGTGTGATTGTCACTGCTCATGGT 120
 DB 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
 QY 121 GTGGGTCTCCAAACGAGCGCCAGCTGATCAGATGTCACCGATGTCACCGAGCGATGCCA 180
 DB 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 QY 181 ACCTGGCCGCTAAACATGGGTGTGGCTGACACCGCGNGGTGGTGAAGTAAATCGCTATCTC 240
 DB 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
 QY 241 GGAACATTACCAACCGAGCTCTGTGTAGCCGAGGATTTGTCGCGAAAAAGGGCTGCCCG 300
 DB 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 QY 301 CAGCTGTGTCAGTCCGATATTGCTGTGATGTCGAGCGGTACCCGTAAACAGCGGAA 360
 DB 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 QY 361 GCCTTCGCCCGCGCTGGCACCTGCTGTGCAATACCGTACATACCCAGGAGATACG 420
 DB 121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 QY 421 TCAGTCCCGATCCGTTATTATTCCTCTAAACCTGGCGTTTGCACACTGGATAACGCG 480
 DB 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 QY 481 AAGTGACTACCGGATCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTACCGGGCAT 540
 DB 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 QY 541 CGCAACAGCGGTTTCGCAACTGGACCGGTGCTTAATTTCCGCAATCAAACTGTGTC 600
 DB 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 QY 601 CTTAAAGCTGAGAAACAGGACGAAGCTGTTCATTAAACGAGGCAATTACCATCGGAATC 660
 DB 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 QY 661 AAGTGAGCCCGCAATGCTCTATTAAACGGTTCGGTAAAGCTCGGATCAATCTGTCAGC 720
 DB 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 QY 721 GAGATATTTCTCTGCAACAGCACAGGAAATCGCGAGCGCGGGTGGGGAAGGATCAC 780
 DB 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260

QY 781 GATTACACAGTGGAACACCTTGCTAAGCTTTTCATAACGCGCAATTTTATTGCTACAA 840
 DB 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
 QY 841 CGCACCCAGAGGTTGCCCGCAGCCGCGCCAGCCCGCTTATTGATTTGATCATGGCAGCG 900
 DB 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 QY 901 TTGACGCCCATCCACCGCAAAACAGGCGTATGGTGTGACATATACCCACTTCAGTACTG 960
 DB 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 QY 961 TTTATTGCGGACACGATATACTTCTGGCAAAATCTCGCGCGCCACTTGGAGCTCAACTGG 1020
 DB 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 QY 1021 ACGTTTCCCGGTGAGCCGATACACCGCCGAGGTGGTGAAGTGGTGTGAACTGGTGG 1080
 DB 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 QY 1081 CGTCGCTAAGCGATTAACGCCAGTTCAGATTTCAGTTTCGCTGCTTTCAGACTTTACAG 1140
 DB 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 QY 1141 CAGATCGGTGATAAAACGCGCTGTCTATTAAATACGCGCCGCGAGAGGTGAACTGACC 1200
 DB 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 QY 1201 CTGGCAGGATGTGAAGAGCGAAATGCGAGGCGATGTTTCGTTGGCAGGTTTACGCAA 1260
 DB 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 ATCGTGAATGAACGACGATACCGCGCTGCAGTTTGAGATCTCATCCATCCATCCATCAC 1320
 DB 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440

RESULT 3
 AAE15806
 ID AAE15806 standard; protein; 440 AA.
 XX
 AC AAE15806;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Escherichia coli B phytase protein.
 XX
 KW Bacterial phytase; K12 appA phytase; protease stability; anabolic;
 KW gastrointestinal; nutritional value; feed treatment process; therapy;
 KW thermal tolerance; growth performance; alcoholic drink; biopulping;
 KW non-alcoholic drink; biobleaching; B phytase.
 XX
 OS Escherichia coli.
 XX
 PN WO200109333-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 24-MAY-2001; 2001WO-US017118.
 XX
 PR 25-MAY-2000; 2000US-00580515.
 XX
 PA (DIVE-) DIVERSA CORP.
 XX
 PI Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'donoghue B;
 XX
 DR WPI; 2002-083108/11.
 DR N-PSDB; AAD25460.
 XX
 PT New bacterial phytase for e.g. improving the nutritional value of phytate
 PT -containing foodstuffs and subsequently improving the growth performance
 PT of an organism that consumes it, or in treating animal digestive systems.
 XX
 PS Claim 1; Fig 1; 170pp; English.

XX The patent discloses recombinant bacterial phytase from *Escherichia coli*
 CC K12 appA phytase. The enzyme has phytase activity and improved thermal
 CC tolerance when compared with wild-type phytase. It has improved protease
 CC stability at low pH. The recombinant phytase is useful for improving the
 CC nutritional value of phytate-containing feedstuffs and subsequently
 CC improving the growth performance of an organism that consumes it, in
 CC treating animal digestive systems, in feed treatment processes and for in
 CC vitro purposes related to research, discovery and development. They are
 CC also used for generating recombinant digestive system life forms, for
 CC producing or manufacturing alcoholic and non-alcoholic drinks based on
 CC the use of moulds, grains and/or plants, in biopulping and bio-bleaching
 CC where a reduction in the use of environmentally harmful chemicals that
 CC are traditionally used in the pulp and paper industry is desired and in
 CC the reduction or possible elimination of the need for mineral
 CC supplements, enzymes or therapeutic drugs for animals from the daily feed
 CC thus increasing the amount calories and nutrients present in the feed.
 CC The present sequence is *E. coli* B phytase protein
 XX

SQ Sequence 440 AA;

Alignment Scores:

Pred. No.: 8,77e-230 Length: 440
 Score: 2302.00 Matches: 440
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.00% Indels: 0
 DB: 5 Gaps: 0

US-10-034-985-1 (1-1323) x AAE15806 (1-440)

QY 1 ATCAAGCGATCTTAATCCCATTTTATCTCTGATTCGTTAAACCCGCAATCTGCA 60
 DB 1 MetyleAlaLeulleProPheLeuSerLeulleProLeuThrProGlnSerAla 20
 QY 61 TTCGCTCAGAGTGAAGCTGAAGTCTGAAGTGTGTGATTTGTGCTGCTCATGTT 120
 DB 21 PheAlaGlnSerGluProGluLeuLeuValValIleValSerArgHisGly 40
 QY 121 GTGCTGCTCAACCAAGCCAGCACTGATGAGAGTGTACCCAGAGCGATGCCA 180
 DB 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 QY 181 ACCTGGCGGTAAACCTGGTGTGCTGACACCCGNGTGTGTGAGCTTAATCGCTATCTC 240
 DB 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyLeuLeuIleAlaTrpLeu 80
 QY 241 GGACATTACCAAGCCAGCGCTCTGTAGCCAGCGATTGTCGCGAAAAAGGCGTCCCG 300
 DB 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 QY 301 CAGTGTGTCAGTGGGATATTGTCGATGTCGACAGCGTACCGGTAAACAGCGGAA 360
 DB 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 QY 361 GCCTTCGCGCGCGGTGGACCTGCTGTCGATGTCATACCTACCCAGGAGATACG 420
 DB 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 QY 421 TCAGTGTCCCGATCGTTATTTATCTCTCTAAACCTGGCGTTTCCCAACTGGATAACGCG 480
 DB 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 QY 481 AAGTGTACTGACGCGATCTCTCAGAGGCGAGGAGGTCAATTCGTGATCTTTACCGGCAT 540
 DB 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 QY 541 CGGCAACCGCGTTTCGCACTGGACGGGTCTTAATTTCCGCAATCAAACTGTGCG 600
 DB 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuLys 200
 QY 601 CTTAAACGTGAGAACAGCAAGCAAGCTGTTTCATTACGCGAGGCAATTCACCATCGAATC 660

DB 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 QY 661 AAGTTCAGCGCGACAAATGCTCATTAACCGGTGCGTAAGCCTCGCATCAATGCTGACG 720
 DB 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 QY 721 GAGATATTTCTCTCTGCAACAAGCACAGGGAATCCCGAGCCGGGTGGGGAAGATCAC 780
 DB 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 QY 781 GATTTCACAGTGGAAACCTTGTGTAAGTTTCATTAAGCGCAATTTATTGCTGTACAA 840
 DB 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
 QY 841 CGCACCCAGAGGTTCGCCGACCGCCGCCACCCGTTATTGATTGATTCATGATGCGAGCG 900
 DB 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 QY 901 TTGACGCCCATCCACCCGCAAAACAGGCGGTATGCTGTGACATTATCCCACTTCAGTACTG 960
 DB 301 LeuThrProHisProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
 QY 961 TTTATTGCGCGACACGATACTAATCTGGCAAAATCTCGCGCGCGCATGAGACTCAACTGG 1020
 DB 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 QY 1021 ACGTTTCCCGGTGAGCGGATACACCGCCGAGGTGGTGAACCTGTTGTTGACCGCTGG 1080
 DB 341 ThrLeuProGlyGlnProAspAsnThrProGlyGlyGluLeuValPheGluArgTrp 360
 QY 1081 CGTGGCTTAAGCGATAACAGCCAGTGTGATTCAGTTCGCTGCTTCCAGACTTTTACAG 1140
 DB 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 QY 1141 CAGATCGGTGATAAAAACGCGCTGTGTCATTAAATACGCGCCCGAGAGGTGAAACTGACC 1200
 DB 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProGlyGluValLysLeuThr 400
 QY 1201 CTGCGAGGATGCAAGAGCGAAATGCGAGGCGATGTTGTTGTCGAGGTTTTCAGCAA 1260
 DB 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 ATCGTGAATGAAGCACGATACCGCGGTGTCAGTTTGTGATCTCATCCATCACCATCAC 1320
 DB 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440

RESULT 4

ADA19446
 ID ADA19446 standard; protein; 440 AA.

AC ADA19446;

XX 20-NOV-2003 (first entry)

DE *E. coli* B phytase.

XX Phytase; enzyme; phytate; appA gene; animal feed; inorganic phosphate;
 KW digestion enhancement; transgenic; thermal tolerance; protease stability.
 XX *Escherichia coli* B.

OS Key Location/Qualifiers

XX Misc-difference 72 /note= "Encoded by CGN"

FT US2002136754-A1.

XX 26-SEP-2002.

XX 24-MAY-2001; 2001US-00866379.

XX 13-AUG-1997; 97US-00910798.

PR 01-MAR-1999; 99US-00259214.

PR	13-APR-1999;	99US-00291931.	QY	121	GTGCGTGTCCAAACCAAGCCAGCAACTGATGTCAGAGATGTCACCCAGACGCGATGGCCA	180
PR	25-MAY-1999;	99US-00318528.	DB	41	ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro	60
XX	25-MAY-2000;	2000US-00580515.				
PA	(SHOR/) SHORT J M.		QY	181	ACCTGGCCGCTAAACCTGGGTTGGCTGACACCCCGNGGTGGTGGAGCTAAATCGCTATCTC	240
PA	(KRET/) KRETZ K A.		DB	61	ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyLeuLeuIleAlaTyLeu	80
PA	(GRAY/) GRAY K A.					
PA	(BART/) BARTON N R.		QY	241	GGACATTACCAACGCCAGCGTCTGGTAGCCGCGATTCTGCGGCAAAAAGGCGTCCCG	300
PA	(GARR/) GARRETT J B.		DB	81	GlyHisTyGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro	100
PA	(DONO/) O' DONOGHUE E.					
PA	(MATH/) MATHUR E J.		QY	301	CAGTCTGGTCCAGTCCGGATTATTGTCGACGAGCGTACCCGCTAAACAGGCGAA	360
PI	Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB;		DB	101	GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu	120
XX	O' Donoghue E, Mathur EJ;					
XX	WPI; 2003-040002/03.		QY	361	GCCTTCGCGCGCGCTGGCACCTGACTGTGCAATAAACCGTACATACCCAGGAGATACG	420
XX	Isolated Escherichia coli polynucleotide encoding a modified phytase		DB	121	AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
PT	enzyme, useful in the production of animal feed, for improving the		QY	421	TCAGTCCCGATCCGTTATTATCTCTTAACTGCGGCTTTGCCAACTGGATAACGCG	480
PT	nutritional value of phytate-containing foodstuff and for enhancing		DB	141	SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla	160
PT	digestion in humans and animals.					
XX	Claim 1; Fig 1; 62pp; English.		QY	481	AACGTGACTGACCGGATCTCTCAGCAGGCGAGGGGTCAATTCGTGACTTTACCGGCAT	540
XX	The invention relates to an isolated Escherichia coli polynucleotide		DB	161	AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis	180
CC	encoding a phytase enzyme appearing as ADA19450 and having amino acids		QY	541	CGCAACACGGCGTTTCGCGAACCTGGAACGGGTCTTAATTTCCGCAATCAAACTCTGCG	600
CC	modification from W68E, Q84W, A95P, K97C, S168E, R180Y, N226C or Y277D.		DB	181	ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys	200
CC	Also included the E. coli appA gene ADA19449 (or an oligonucleotide		QY	601	CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAAACGACGACATTACCATCGAACTC	660
CC	derived from it) or its mutant sequence ADA19452, expression vectors,		DB	201	LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu	220
CC	host cells, a method of improving nutritional value of a phytate-		QY	661	AAGTGAGCCGCCAGCAATGCTCTATTAAACGGTGGGTAGCTCGCATCAATGCTACAG	720
CC	containing foodstuff by contacting the phytate-containing foodstuff with		DB	221	LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr	240
CC	a pure phytase enzyme mutated phytase (where the phytase enzyme catalyzes		QY	721	GAGATATTTCTCTGCAACAGCACAGGAAATCCGGAGCCGGGTGGGAGGATCACC	780
CC	the liberation of inorganic phosphate from the phytate in the phytate-		DB	241	GluIlePheLeuLeuGlnGlnAlaGlnGlyWecProGluProGlyTrpGlyArgIleThr	260
CC	containing foodstuff), a method to produce an animal feed containing a		QY	781	GATTTCACACAGTGGAAACACTTGTCTGTAAGTTTGCATAACGCGCAATTTATTTGCTACAA	840
CC	microbial phytase (comprising culturing the plant cell, plant part or		DB	261	AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyLeuLeuGln	280
CC	plant under conditions where the nucleotide sequence is expressed and		QY	841	CGACGCGCAGAGTTCGCCGACGCGCGCCACCCCGTTATTGATTTGATCATGGCAGCG	900
CC	converting the plant cells, plants or plant into a composition for animal		DB	281	ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla	300
CC	feed), a feed composition for animals (comprising the plant seeds, plant		QY	901	TTGACGCCCATCCACCGCAAAAACAGGCGGTATGGTGCATATTACCCACTTCAGTACTG	960
CC	cells, plant parts or plants in admixture with a phytate-containing		DB	301	LeuThrProHisProProGlnLysGlnAlaTyGlyValThrLeuProThrSerValLeu	320
CC	foodstuff), a method to treat a human or an animal able to benefit from		QY	961	TTTATTGCGCGACACGATATAATCTGGCAAAATCTCGCGCGCGCAGCTCGAGCTCAACTGG	1020
CC	digestive enhancement by the activity of an exogenous phytase enzymes		DB	321	PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp	340
CC	comprising administering to the human or animal the plant seed, plant		QY	1021	ACGCTTCCCGGTGACGCGGATAAACGCCCGCGAGTGGTGAACCTGGTTTGAACGCTGG	1080
CC	cells, plant parts or plants of a transgenic plant which is modified to		DB	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp	360
CC	contain an expression system which expresses a nucleotide sequence		QY	1081	CGTGGGTAGCGGATACACCGCAGCTGGATTTCAGTTTCGTTGCTTCCAGACTTTACAG	1140
CC	encoding a phytase enzyme, a transgenic non-human organism whose genome		DB	361	ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln	380
CC	comprising a heterologous nucleic acid sequence encoding a polypeptide		QY	1141	CAGATGGGTGATAAAACCGCGCTGTCTATTAAATACGCCCGCGGAGAGGTGAACTGACC	1200
CC	having phytase activity. The phytase enzyme is useful for improving the		DB	381	GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr	400
CC	nutritional value of phytate-containing foodstuff, in the production of		QY	1201	CTGCGAGGATGTGAAGACGCAAAATGCGCAGGGCATGTGTTGCTTGGCAGGTTTTACGCA	1260
CC	animal feed and for enhancing digestion in humans and animals. The		DB			
CC	invented method improves thermal tolerance and protease stability. It					
CC	also improves the feeding value of phytate rich ingredients. The present					
CC	sequence represents E. coli B wild-type phytase.					
XX	Sequence 440 AA;					
SQ						
Alignment Scores:						
Pred. No.:	8.77e-230	Length:	440			
Score:	2302.00	Matches:	440			
Percent Similarity:	100.00%	Conservative:	0			
Best Local Similarity:	100.00%	Mismatches:	0			
Query Match:	98.00%	Indels:	0			
DB:	6	Gaps:	0			
US-10-034-985-1 (1-1323) x ADA19446 (1-440)						
QY	1	ATGAAGCGATCTTAATCCCAATTTTATCTCTTGTGATTCGGTTAACCCCGCAATCTGCA	60			
DB	1	MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla	20			
QY	61	TTCCCTCAGATGAGCCGAGCTAGCTGAAAGTGTGGTATGTGTCAGTCGTCATGGT	120			
DB	21	PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly	40			

Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 ATGCTGAATGAGCACCACATACCGGCTGAGTTGAGATCTATCACCATCACATCAC 1320
 Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440

RESULT 5
 ADOS0292
 ID ADOS0292 standard; protein; 440 AA.
 XX AC ADOS0292;
 XX DT 29-JUL-2004 (first entry)
 DE Escherichia coli B phytase enzyme.
 XX KW Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
 KW KW phytate; animal feed; fish feed; dough; baking; enzyme.
 XX OS Escherichia coli; B.
 XX FH Key Location/Qualifiers
 FT Misc-difference 72 /note= "Encoded by CGN"
 XX PN US2004091968-A1.
 XX PD 13-MAY-2004.
 XX PF 20-JUN-2003; 2003US-00601319.
 XX PR 13-AUG-1997; 97US-00910798.
 PR 01-MAR-1999; 99US-00259214.
 PR 13-APR-1999; 99US-00291931.
 PR 25-MAY-1999; 99US-00318528.
 PR 25-MAY-2000; 2000US-00580515.
 PR 24-MAY-2001; 2001US-00866379.
 XX (SHORT) SHORT J M.
 PA (KRETT) KRETT K.
 PA (BART) BARTON N R.
 PA (GARR) GARRETT J B.
 PA (ODON) O'DONOGHUE E.
 PA (MATH) MATH E J.
 XX Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E,
 PI Mather EU;
 XX WPI: 2004-374952/35.
 DR N-PSDB; ADOS0291.
 XX
 CC Producing phytase, involves providing nucleic acid derived from bacteria
 FT encoding polypeptide having phytase activity, and expressing nucleic acid
 PT in yeast.
 XX
 PS Claim 3; SEQ ID NO 2; 74pp; English.
 XX
 CC The invention relates to a method for producing phytase that involves
 CC providing a nucleic acid encoding phytase derived from a bacteria, and
 CC expressing the nucleic acid in a yeast under conditions that allow
 CC expression of the enzyme in the yeast. The invention also relates to
 CC modified phytase enzyme which has improved thermal tolerance and protease
 CC stability at low pH. The phytase enzyme can be used in foodstuffs to
 CC improve the feeding value of phytate rich ingredients, and in diet of
 CC numerous animals including mammals, fowls and fishes, commercially
 CC significant animals such as pigs, goats, laboratory rodents, commercially
 CC significant avian species such as chicken, ducks, doves, parrot, etc.,
 CC commercially farmed fish such as guppy, zebrafish, molly, swordtail,
 CC etc., in dough making and baking, in dietary aids for animals. The method
 CC provides easy manufacture of the active ingredient loaded biocompatible
 CC composition, higher yields and loading efficiency. The phytase

CC incorporated in to the dietary aid is safe for animals. The present
 CC sequence is Escherichia coli B phytase enzyme.
 XX
 SQ Sequence 440 AA;
 Alignment Scores:
 Pred. No.: 8,77e-230 Length: 440
 Score: 2302.00 Matches: 440
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 98.00% Indels: 0
 DB: 8 Gaps: 0
 US-10-034-985-1 (1-1323) x ADOS0292 (1-440)
 QY 1 ATGAAAGCATCTTAATCCCATTTTATCTCTTCTGATTCGGTTAAACCCCGCAATCTGCA 60
 Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 QY 61 TTCGCTCAGAGTCAGCCGAGCTGAAGCTGGAAGTGGTGAATTCAGTCGTCATGGT 120
 Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
 QY 121 GTGCTGTCTCCAAACCAAGGCCACGCAACTGATCAGGATGTACCCACGATGSCCA 180
 Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 QY 181 ACCTGGCCGTAATAACTGGGTGGCTGACACCGCGGGTGGTGGTGAATTCGCTATCTC 240
 Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyLeuIleAlaTrLeu 80
 QY 241 GGACATTACCAACGCCAGCGCTCTGGTAGCCGCGATTGCTGCGGAAAAAGGGCTGCCG 300
 Db 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 QY 301 CAGTCTGGTCAGTCGCGATTATTGCTGATGTCCAGAGCGGTACCGTAAAAACAGCGCAA 360
 Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 QY 361 GCCTTCGCCCGCGGCTGGCAGCTGCTGCAATTAACCGTACATACCCAGGCGAGATACG 420
 Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 QY 421 TCCAGTCCCGATCCGTTATTATCTCTAAACCTGGGCTTTCGCAACTGGATAACGCG 480
 Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAla 160
 QY 481 AACGTGACTGACCGATCTCTCAGCGGAGGAGGTCAATTGCTGACTTTACCGGCGAT 540
 Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 QY 541 CGGCAACCGCGTTTCGCAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 600
 Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuLys 200
 QY 601 CTTAAACGTCAGAAAAACAGACGAAAGCTGTTCATTAAACGAGGCGATTTACCATCGAACTC 660
 Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 QY 661 AAGTGAGCGCCGACCAATGTCTCATTTAAACGGTGGGTGAAGCCTCGCATCAATCTGACG 720
 Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 QY 721 GAGATATTTCTCTGCAACAGCACAGGGAATCCCGAGCCGGGTGGGGAGGATCACC 780
 Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 QY 781 GATTTCACACAGTGAACACCTTTGCTTAAGTTTCATAACGCGCAATTTTATTTGCTACAA 840
 Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
 QY 841 CGCACGCCAGAGGTTCGCCGACGCCGCCCGCTTATGATTTGATTCATGGCAGCG 900

Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 QY 901 TTGACGGCCCATCCACCGCAAAACAGCGGTATGGTGTGATACCATTCAGTACTG 960
 Db 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 QY 961 TTTATTGCGGACAGCATACTAATCTGGCAAAATCTCGCGCGCACTGGAGCTCAACTGG 1020
 Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsnTrp 340
 QY 1021 ACGCTTCCCGGTACGCGGATTAACACCGCCGAGGTGGTGAACCTGGTGTGAACTGG 1080
 Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 QY 1081 CGTGGCTAAGCGATACACCGCATCGGATTCAGTTTCGTGTGCTTCCAGACTTTACAG 1140
 Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 QY 1141 CAGATGGCTATAAAGCGCGCTGTCAATTAATACGCGCGCGGAGAGGTGAACCTGACC 1200
 Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 QY 1201 CTGCAGGATGTGAAGAGCGAAATCGCAGCGCATGTGTTCTGGCAGGTGTTTACGCAA 1260
 Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 ATCGTGAATGAAGCAGCATACCGCGGTGCGAGTTTGAGATCTCATCACCATCATCAC 1320
 Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440

RESULT 6
 ADO50304
 ID ADO50304 standard; protein; 440 AA.
 AC ADO50304;
 XX
 XX
 DT 29-JUL-2004 (first entry)
 DE
 DE Escherichia coli W phytase 875PH2 mutant enzyme.
 XX
 KW Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
 KW phytate; animal feed; fish feed; dough; baking; mutant; mutin.
 XX Escherichia coli.
 OS
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 72 /note= "Encoded by CGN"
 FT Misc-difference 160 /note= "E. coli B phytase Ala replaced with Ser"
 FT Misc-difference 176 /note= "E. coli B phytase Asp replaced with Gly"
 FT Misc-difference 298 /note= "E. coli B phytase Met replaced with Lys"
 FT Misc-difference 299 /note= "E. coli B phytase Ala replaced with Thr"
 XX
 XX US2004091968-A1.
 XX
 PD 13-MAY-2004.
 XX
 XX 20-JUN-2003; 2003US-00601319.
 XX
 PR 13-AUG-1997; 97US-00910798.
 PR 01-MAR-1999; 99US-00259214.
 PR 13-APR-1999; 99US-00291931.
 PR 25-MAY-1999; 99US-00318528.
 PR 25-MAY-2000; 2000US-00580515.
 PR 24-MAY-2001; 2001US-00866379.
 XX
 XX (SHORT) SHORT J M.
 PA (KRETZ) KRETZ K.
 PA (GRAY) GRAY K A.

PA (BART/) BARTON N R.
 PA (GARR/) GARRETT J B.
 PA (ODON/) O'DONOGHUE E.
 PA (MATH/) MATHER E J.
 XX
 PI Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
 Mather EJ;
 XX
 XX WPI; 2004-374952/35.
 DR N-PSDB; ADO50291, ADO50303.
 XX
 XX Producing phytase, involves providing nucleic acid derived from bacteria
 encoding polypeptide having phytase activity, and expressing nucleic acid
 in yeast.
 PS
 PS Disclosure; Page: 74pp; English.
 CC The invention relates to a method for producing phytase that involves
 providing a nucleic acid encoding phytase derived from a bacteria, and
 expressing the nucleic acid in a yeast under conditions that allow
 expression of the enzyme in the yeast. The invention also relates to
 modified phytase enzyme which has improved thermal tolerance and protease
 stability at low pH. The phytase enzyme can be used in foodstuffs to
 improve the feeding value of phytate rich ingredients, and in diet of
 numerous animals including mammals, fowls and fishes, commercially
 significant mammals such as pigs, goats, laboratory rodents, commercially
 significant avian species such as chicken, ducks, doves, parrot, etc.,
 commercially farmed fish such as guppy, zebrafish, molly, swordtail,
 etc., in dough making and baking, in dietary aids for animals. The method
 provides easy manufacture of the active ingredient loaded biocompatible
 composition, higher yields and loading efficiency. The phytase
 incorporated in to the dietary aid is safe for animals. The present
 sequence is Escherichia coli W phytase mutant enzyme. Note: The present
 sequence is not shown in the specification but has been derived from
 CC Escherichia coli B phytase ADO50292.
 XX
 SQ Sequence 440 AA;

Alignment Scores:
 Pred. No.: 1,06e-227 Length: 440
 Score: 2282.00 Matches: 436
 Percent Similarity: 99.32% Conservative: 1
 Best Local Similarity: 99.09% Mismatches: 3
 Query Match: 97.15% Indels: 0
 DB: 8 Gaps: 0

US-10-034-985-1 (1-1323) x ADO50304 (1-440)

QY 1 ATGAAGCGATCTTAATCCATTTTATCTCTGATTCGTTACCCCGCAATCTGCA 60
 Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 QY 61 TTCGCTCAGAGTCAGCGCGGAGCTGAAGCTGGAAGTGGTGAATGTGTCAGTCGTAGGT 120
 Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
 QY 121 GTCGCTGCTCCAAACCAAGCCAGCAACTGATCAGAGATGTCACCCAGACGATGSCCA 180
 Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 QY 181 ACCTGCGCGGTAAACTGGGTTGGCTGACACCCGCGGAGTGGTGAAGTAACTGCTATCTC 240
 Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
 QY 241 GGACATTACCAACCGCAGCGCTGTGGTAGCCGAGGATTCGTCGCGCAAAAGGGCTGCCG 300
 Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 QY 301 CAGTCTGGTCAGTCGCGATTTATGCTGATGTCAGCAGCGGTACCCGCTAAACAGCGCAA 360
 Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 QY 361 GCCTTCGCGCGCGGCTGGCACCTGACTGTGCAATAACCGTATACATACCCAGGAGATACG 420

Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 Qy 421 TCAGTCCCGATCGTATTATTAATCTCTAAACAGCGCTTGCACAACTGGATAACGCG 480
 Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnSer 160
 Qy 481 AACGTGACTCAGCGATCTCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTTACCGGCAT 540
 Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaGlyPheThrGlyHis 180
 Qy 541 CGGCAACCGCGTTTCGGAACCTGGAACCGGTGCTTAATTTCCGCAATCAAACTGTGTC 600
 Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 Qy 601 CTTAAACGTCAGAACACAGGACGAAGCTGTTTCATTACGAGCAGCATTTACCATCGGAATC 660
 Db 201 LeuLysArgGlnLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 Qy 661 AAGGTGAGCCCGACAAATGTTCTATTAAACCGGTGCGGTAAAGCCTCGCATCAATGTCAGC 720
 Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 Qy 721 GAGATATTCTCTGCAACAGCACAGGGAATGCGGAGCGCGGGTGGGAAGATCACC 780
 Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 Qy 781 GATTACACAGGTGAACACTTCTGTAAGTTTCATACACGCGCAATTTTATTGTTGCTACAA 840
 Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
 Qy 841 CGCACGCGCAGAGGTGCGCCGCGCAGCGCCGCCCGCTTATTGGATTTGATCATGCGAGCG 900
 Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
 Qy 901 TTGACGCGCCCATCACCGCAAAACAGCGGTATGGTGTGACATTACCATTCAGTACTG 960
 Db 301 LeuThrProHisProGlnLysGlnAlaTyxGlyValThrLeuProThrSerValLeu 320
 Qy 961 TTTATTGCGGACGACGATCTTAATCTGCAAAATCTCGCGCGCACCTGGAGCTCAACTGG 1020
 Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 Qy 1021 ACGTTCCTCCGCTCAGCGGATAACACGCGCCGCGGAGGTGAGTGGTGTGAACTGGTGG 1080
 Db 341 ThrLeuProGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 Qy 1081 CGTCGGCTAAGCGATAACAGCCAGTGGATTCAGGTTTCGCTTCGAGCTTTACAG 1140
 Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 Qy 1141 CAGATGGTGTATAAACCGCGCTGTCTATTAAATACGCGCGCGAGAGGTGAAACTGAC 1200
 Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 Qy 1201 CTGCGAGGATGTGAAGACGAAATTCGCGAGGATGTGTTTCGTTGGCAGGTTTTACGCAA 1260
 Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Qy 1261 ATGTTGAATGAGCAGCAGCATACCGGTGAGTTTGAGTCTCATTCACATCCATCATCAC 1320
 Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
 RESULT 7
 ID ADO50302 standard; protein; 440 AA.
 AC AC
 XX AC
 XX AC
 DT 29-JUL-2004 (first entry)
 XX Kangaroo rat Escherichia coli phytase 872PH1 mutant enzyme.
 DE
 XX

KW Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
 KW phytate; animal feed; fish feed; dough; baking; mutant; mutein.
 OS Escherichia coli.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 72 /note= "Encoded by CGN"
 FT Misc-difference 176 /note= "E. coli B phytase Asp replaced with Gly"
 FT Misc-difference 298 /note= "E. coli B phytase Met replaced with Lys"
 FT Misc-difference 299 /note= "E. coli B phytase Ala replaced with Thr"
 FT Misc-difference 312 /note= "E. coli B phytase Gly replaced with Ser"
 XX
 PN US2004091968-A1.
 XX
 XX 13-MAY-2004.
 XX
 PF 20-JUN-2003; 2003US-00601319.
 XX
 XX 13-AUG-1997; 97US-00910798.
 PR 01-MAR-1999; 99US-00259214.
 PR 13-APR-1999; 99US-00291931.
 PR 25-MAY-1999; 99US-00318528.
 PR 25-MAY-2000; 2000US-00580515.
 PR 24-MAY-2001; 2001US-00866379.
 XX
 PA (SHOR/) SHORT J M.
 PA (KRET/) KRETZ K.
 PA (GRAY/) GRAY K A.
 PA (BART/) BARTON N R.
 PA (GARR/) GARRETT J B.
 PA (ODON/) O'DONOGHUE E.
 PA (MATH/) MATHER E J.
 XX
 PI Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
 PI Mather EJ;
 XX
 DR WPI; 2004-374952/35.
 DR N-PSDB; ADO50301.
 XX
 PT Producing phytase, involves providing nucleic acid derived from bacteria
 PT encoding polypeptide having phytase activity, and expressing nucleic acid
 PT in yeast.
 XX
 PS Disclosure; Page; 74pp; English.
 XX
 CC The invention relates to a method for producing phytase that involves
 CC providing a nucleic acid encoding phytase derived from a bacteria, and
 CC expressing the nucleic acid in a yeast under conditions that allow
 CC expression of the enzyme in the yeast. The invention also relates to
 CC modified phytase enzyme which has improved thermal tolerance and protease
 CC stability at low pH. The phytase enzyme can be used in foodstuffs to
 CC improve the feeding value of phytate rich ingredients, and in diet of
 CC numerous animals including mammals, fowls and fishes, commercially
 CC significant mammals such as pigs, goats, laboratory rodents, commercially
 CC significant avian species such as chicken, ducks, doves, parrot, etc.,
 CC commercially farmed fish such as guppy, zebrafish, molly, swordtail,
 CC etc., in dough making and baking, in dietary aids for animals. The method
 CC provides easy manufacture of the active ingredient loaded biocompatible
 CC composition, higher yields and loading efficiency. The phytase
 CC incorporated in to the dietary aid is safe for animals. The present
 CC sequence is Kangaroo rat Escherichia coli phytase mutant enzyme. Note:
 CC The present sequence is not shown in the specification but has been
 CC derived from Escherichia coli B phytase sequence ADO50292.
 XX
 SQ Sequence 440 AA;
 Alignment Scores: 2.17e-227 Length: 440
 Pred. No.:

Score:	2279.00	Matches:	436
Percent Similarity:	99.09%	Conservative:	0
Best Local Similarity:	99.09%	Mismatches:	4
Query Match:	97.02%	Indels:	0
DB:	8	Gaps:	0
US-10-034-985-1 (1-1323) x AD050302 (1-440)			
QY	1	ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGTTAAACCCCGCAATCTGCA	60
DB	1	MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla	20
QY	61	TTGCTCTCAGAGTCAGCGGAGCTGAAGCTGCGAAAGTGTGTGATTTGTCAGTCTGATGTT	120
DB	21	PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly	40
QY	121	GTCCGTGCTCCAAACGAGGCCAGCAATGATCAGGATGTCAACCCAGACGATGGCCA	180
DB	41	ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro	60
QY	181	ACCTGCGCGTAAACTGGGTTCGCTGACACCCGNGGTGGTGAGCTAATCGCCTATCTC	240
DB	61	ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu	80
QY	241	GGACATTACCAACGCCAGCGTCTGGTAGCCGACGATTTGTCGGCGAAAAGGGCTGCCCG	300
DB	81	GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro	100
QY	301	CAGTCTGCTCAGCTCCGATTATTGCTGACGAGCGGTACCGGTAACCGTAACAGCGGAA	360
DB	101	GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu	120
QY	361	GCCTTCGCGCGCGGTGGCACCTGTGCAATTAACCGTACATACCCAGCGAGATACG	420
DB	121	AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
QY	421	TCAGTCCCGATCCGTTATTATTCCTTAAACCTGGCGTTTGGCAACTGGATAACGCG	480
DB	141	SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla	160
QY	481	AACGTGACTGACCGGATCTCCACGAGCGGAGGAGGTCAATTCGCTGACTTTACCGGCAT	540
DB	161	AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaGlyPheThrGlyHis	180
QY	541	CGCAAAACGCGGTTCCGCAACTGGAACGGGTCTTAATTTCCGCAATCAAACTTGTC	600
DB	181	ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys	200
QY	601	CTTAAACGTGAGAAACAGGACGAAGCTGTTCATTAAACGAGGATTCACATCGGAATC	660
DB	201	LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu	220
QY	661	AGGTGAGCGCCGACATGCTCATTAACGGTGGGTAGCCTCCGATCAATGCTGACG	720
DB	221	LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr	240
QY	721	GAGATATTTCTCTGCAACACAGGAAATCCGAGCCGGGTGGGGAAGGATCAACC	780
DB	241	GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr	260
QY	781	GATTACACAGTGGAAACACTGCTTAAGTTGTCATAACCGCAATTTTATTTGCTACAA	840
DB	261	AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln	280
QY	841	CGACGCCGAGGTGGCCCGCAGCGGCCACCCGTTATGATTTGATCATGGGAGCG	900
DB	281	ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla	300
QY	901	TTGACGCCCATCCACCGCAAAACAGCGCGTATGGTGTGACATTACCCTTCAGTACTG	960
DB	301	LeuThrProHisProProGlnLysGlnAlaTyrSerValThrLeuProThrSerValLeu	320
QY	961	TTTATTGCGGACACGATCACTAATCTGGCAAAATCTCGCGCGCAGCTGGAGCTCAACTGG	1020

Db	321	PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp	340
QY	1021	ACGCTTCCCGGTACCGGATAACCGCCGAGGTGGTGAACCTGGTGTGAACTGG	1080
Db	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp	360
QY	1081	CGTCGGCTAAGCGATAAACAGCCAGTGGATTACAGTTTCCTGCTTCAGACTTTACAG	1140
Db	361	ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln	380
QY	1141	CAGATCGGTGATAAACCGCGTGTCTATTAAATACGCCGCCGAGAGGTGAACCTGACC	1200
Db	381	GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr	400
QY	1201	CTGGCAGGATGTAAGAGCGAAATGCGAGGCGATGTTCGTTGGCAGGTTTACGCCAA	1260
Db	401	LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln	420
QY	1261	ATCGTGAATGAAGCAGCGATACCGGCGTGCAGTTTTCAGATCTCATCACCATCCATCAC	1320
Db	421	IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis	440
RESULT 8			
AAB36257			
ID	AAB36257 standard; protein; 432 AA.		
AC	AAB36257;		
XX			
DT	12-SEP-2003 (revised)		
DT	20-FEB-2001 (first entry)		
XX			
DE	Lama2/APPA plasmid translated sequence.		
DE	Transgenic animal; salivary protein; phytase; phosphorus; animal growth;		
KW	environmental pollution; pig.		
KW			
OS	Mus musculus.		
OS	Escherichia coli.		
OS	Chimeric.		
XX			
PN	W0200064247-A1.		
XX			
PD	02-NOV-2000.		
XX			
PF	20-APR-2000; 2000WO-CA000430.		
XX			
PR	23-APR-1999; 99US-0130508P.		
XX			
PA	(UYGU-) UNIV GUELPH.		
XX			
PI	Forsberg CW, Golovan S, Phillips JP;		
XX			
DR	WPI; 2000-687245/67.		
DR	N-PSDB; AAC68294.		
XX			
PT	Transgenic non-human animal for gastrointestinal tract specific		
PT	expression of a protein, preferably phytase, comprises a nucleic acid		
PT	sequence including a heterologous transgene construct encoding the		
PT	protein.		
XX			
PS	Disclosure; Fig 5; 152pp; English.		
XX			
CC	The present invention provides transgenic animals which produce desired		
CC	proteins, in this case pigs which expresses phytase in the salivary		
CC	gland. Low phytase production levels result in phytate in the diet being		
CC	excreted and causing phosphorus contamination in water, as well as		
CC	reducing the growth of animals. The invention provides a number of		
CC	transgenes containing the E. coli APPA phytase coding sequence. (Updated		
CC	on 12-SEP-2003 to standardise OS field)		
XX	Sequence 432 AA;		

Alignment Scores:

Pred. No.: 8,11e-223 Length: 432
 Score: 2235.00 Matches: 0
 Percent Similarity: 99.54% Conservative: 2
 Best Local Similarity: 99.54% Mismatches: 0
 Query Match: 95.15% Indels: 0
 DB: 3 Gaps: 0

US-10-034-985-1 (1-1323) x AAB36257 (1-432)

QY 1 ATGAACCGATCTTAATCCATTTTATCTCTTCTGATCCGTTAAACCCCGCAATCTGCA 60
 Db 1 MettysalaleuileuileProPheLeuSerLeuLeuileProLeuThrProGlnSerAla 20
 QY 61 TTGGCTCAGAGTACCGGAGCTGAAGCTGAAAGTGTGTGATGTGACGTCTGTCATGTT 120
 Db 21 PheAlaGlnSerGluProGluLeuLeuLysLeuGluSerValValIleValSerArgHisGly 40
 QY 121 GTGGTGCTCCAAACCAAGGCCACGCAACTGATGTCAGGATGTCACCCCGACGATGGCCA 180
 Db 41 ValargAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 QY 181 ACCTGGCCGTAATAACCTGGTGTGCTGACACCGCGNGGTGGTGAAGTAACTCGCTATCTC 240
 Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
 QY 241 GGACATTACCAACCGCAGCTCTGGTAGCCGACGGATTGCTGGGAAAAAGGGCTGCGCG 300
 Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaTysGlyCysPro 100
 QY 301 CAGTCTGGTCAGGTCGGGATATTGCTGATGTGCACGAGCGTACCCGTAACACAGGCGAA 360
 Db 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 QY 361 GCCTTCGCCCGCGGTGGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG 420
 Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 QY 421 TCAGTCCCGATCGTTATTTATCTCTAAACCTGGCGTTTCCCACTGGATGATACGGC 480
 Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 QY 481 AACGTGACTGACGCGATCTCTCAGCAGGCGAGGAGGTCAATTGCTGACTTTTACCGGCGAT 540
 Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 QY 541 CGGCAACCGCGGTTCCGCAACCTGGAACCGGCTTAAATTTCCGCAATCAAACTGTGTC 600
 Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 QY 601 CTTAAACGTGAGAAACAGGACGAAGCTGTTCAATACGCGAGCATTTACCATCGAATC 660
 Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 QY 661 AAGTGAGCGCCGACATGCTCTATTAAACCGGTGCGGTAAGCCTCGCATCAATGCTGACG 720
 Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 QY 721 GAGATATTTCTTCGAACAGCACAGGGAATGCGGAGCCGGGTGGGGAAGATCAACC 780
 Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 QY 781 GATTACACAGTGGAAACCTTGCATTAAGTTTCATTAACCGCAATTTTATTTCGTACAA 840
 Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
 QY 841 CGCACGCGCAGAGTGTCCCGCAGCGCCACCCCGTTATTTGATTTGATCATGGCAGCG 900
 Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
 QY 901 TTGACGCCCATCCACCGCAAAACAGCGCGTGTGTGATCATCCACTTCAGTACTG 960
 Db 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320

QY 961 TTTATTCCGGACACGATATACTAATCTGGCAAAATCTCGGGCGCACTGGAGCTCAACTGG 1020
 Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 QY 1021 ACGTTTCCCGGTGAGCCGGATACACCGCCCGAGGTGGTGAACCTGGTGTGTTGAACGCTGG 1080
 Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 QY 1081 CGTCGGCTAAGCAATACACCGCAGTGGATTACGTTTCGTGCTCTTCCAGACTTTACAG 1140
 Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 QY 1141 CAGATCGGTGATAAAGCGCGTCTCATTAATAATACCGCCCGAGAGGTGAACTGAC 1200
 Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 QY 1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCGATGTGTTGCTGGCAGGTTTTACGCAA 1260
 Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 ATCGTGAATGAACGACGATACCGCGCTGCAGTTTG 1296
 Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
 RESULT 9
 AAB36261
 ID AAB36261 standard; protein; 432 AA.
 AC AAB36261;
 XX
 DT 12-SEP-2003 (revised)
 DT 20-FEB-2001 (first entry)
 XX
 DE R15/APPA plasmid translated sequence.
 XX
 KW Transgenic animal; salivary protein; phytase; phosphorus; animal growth; environmental pollution; pig.
 XX
 OS Rattus sp.
 OS Escherichia coli.
 OS Chimeric.
 PN WO200064247-A1.
 XX
 PD 02-NOV-2000.
 XX
 PF 20-APR-2000; 2000WO-CA000430.
 XX
 PR 23-APR-1999; 99US-0130508P.
 XX
 PA (UYGU-) UNIV GUELPH.
 XX
 PI Forsberg CW, Golovan S, Phillips JP;
 XX
 DR WPI: 2000-687245/67.
 DR N-PSDB; AAC68298.
 XX
 PT Transgenic non-human animal for gastrointestinal tract specific expression of a protein, preferably phytase, comprises a nucleic acid sequence including a heterologous transgene construct encoding the protein.
 PS Disclosure; Fig 21; 152pp; English.
 CC
 CC The present invention provides transgenic animals which produce desired proteins, in this case pigs which express phytase in the salivary gland. Low phytase production levels result in phytate in the diet being excreted and causing phosphorus contamination in water, as well as reducing the growth of animals. The invention provides a number of transgenes containing the E. coli APPA phytase coding sequence. (Updated on 12-SEP-2003 to standardise OS field)
 CC
 XX

Db	301	LeuThrProHisProGlnLysGlnAlaTyGlyValThrLeuProThrSerValLeu	320
Qy	961	TTTATTGCCGGACACGATACTAATCTGGCAAAATCTCGCGCGGCACACTGGAGCTCAACTGG	1020
Db	321	PhelAlaGlyHisapThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp	340
Qy	1021	ACGTTTCCCGGTGAGCCGGATAACACGCCGCCAGGTGGTGAACTGCTGTTGAAACCGCTGG	1080
Db	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp	360
Qy	1081	CGTCGGCTAAGCCATAACAGCCAGTCGATTTCAGGTTTCGCTGCTCTCCAGACTTTACAG	1140
Db	361	ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln	380
Qy	1141	CAGATGGCTGATAAACGCCCGCTGTCTATTAAATACGCCGCCCGAGAGGTGAAACTGACCC	1200
Db	381	GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr	400
Qy	1201	CTGCCAGGATGTGAAGAGCGAAATGCCAGGGCATGTGTTCTGTTGGCAGGTTTTACGCAA	1260
Db	401	LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln	420
Qy	1261	ATCGTGAATCAACAGCAGCATACCGGCGTGCAGTTTG	1296
Db	421	IleValAsnGluAlaArgIleProAlaCysSerLeu	432
RESULT 10			
DB	AAB36262	standard; protein; 432 AA.	
AC	AAB36262;		
XX	12-SEP-2003 (revised)		
DT	20-FEB-2001 (first entry)		
XX	SV40/APPA plasmid translated sequence.		
XX	Transgenic animal; salivary protein; phytase; phosphorus; animal growth; environmental pollution; pig.		
OS	Simian virus 40.		
OS	Escherichia coli.		
XX	Chimeric.		
XX	W0300064247-A1.		
XX	02-NOV-2000.		
XX	20-APR-2000; 2000WO-CA000430.		
XX	23-APR-1999; 99US-0130508P.		
XX	(UYGU-) UNIV GUELPH.		
XX	Forsberg CW, Golovan S, Phillips JP;		
XX	WPI; 2000-687245/67.		
DR	N-PSDB; AAC68299.		
XX	Transgenic non-human animal for gastrointestinal tract specific expression of a protein, preferably phytase, comprises a nucleic acid sequence including a heterologous transgene construct encoding the protein.		
XX	Disclosure; Fig 22; 152pp; English.		
XX	The present invention provides transgenic animals which produce desired proteins, in this case pigs which expresses phytase in the salivary gland. Low phytase production levels result in phytate in the diet being excreted and causing phosphorus contamination in water, as well as reducing the growth of animals. The invention provides a number of transgenes containing the E. coli APPA phytase coding sequence. (Updated		

CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli AppA phytase coding sequence. (Updated
 CC on 12-SEP-2003 to standardise OS field)

XX
 SQ Sequence 432 AA;

Alignment Scores:

Pred. No.: 8.11e-223 Length: 432
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservativity: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 95.15% Indels: 0
 DB: 3 Gaps: 0

US-10-034-985-1 (1-1323) x AAB36259 (1-432)

QY 1 ATGAAGCGATCTTAATCCCAATTTTATCTCTCTGATTCGTTAAACCCCGCAATCTGCA 60
 DB 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 QY 61 TTCCTCAGAGTGAGCGGAGCTGAAGCTGAAAGTGTGTGATTTGTCAGTGTGTCATGGT 120
 DB 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValIleValSerArgHisGly 40
 QY 121 GTGGCTGCTCAACACAGGCCACCACTGATGTCAGGATGTCACCCAGAGCGATGSCCA 180
 DB 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 QY 181 ACCTGGCCGTTAAACCTGGTGTGCTGACACCGGNGTGTGAGCTAATCGCCTATCTC 240
 DB 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaLysLeu 80
 QY 241 GGACATTACCAACGCCAGCGTCTGCTAGCCGACGGATTGCTGGCGAAAAAGGGCTGCCG 300
 DB 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 QY 301 CAGTCTGGTCAGTCCGGATTATTGCTGATGTCGACAGCGTACCCGTAAACAGGCGAA 360
 DB 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 QY 361 GCCTTCGCCGCGGCTGGACCTGACTGTGTCATATACCGTACATACCCAGGACAGATACG 420
 DB 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 QY 421 TCAGTCCCGATCCGTTATTATTAATCTCTAAACCTGCGCTTTCGCAACTGGATAACGCG 480
 DB 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 QY 481 AACGTGACTGACGCGATCCTCAGAGCGGAGGAGGTCAATTGCTGACTTTACCGGCAT 540
 DB 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 QY 541 CGGCAACCGGTTTCGCACTGGAACGGTGTCTTAATTTCCGCAATCAAACTGTGC 600
 DB 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 QY 601 CTTAAACGTCGAGAACAGGAGAAAGCTGTTTCAATTAACGAGGATTAACCATCGGAATC 660
 DB 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 QY 661 AAGGTAGCGCGCAATGCTCTCAATTAAACCGGTGCGGTAGCCTCGCATCATGCTGACG 720
 DB 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 QY 721 GAGTATTTCTCTGCAACAGCACGGATGCGGAGCGGGGTGGGAAGATCACC 780
 DB 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 QY 781 GATTACACACGATGGAACCTCTGCTAAGTTTGATTAACGCGCAATTTTATTTCTACAA 840
 DB 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
 QY 841 CGCAGCCAGAGGTTGCCCGCAGCGCCACCCCGTTATTGGATTGATCATGCGCAGCG 900

DB 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
 QY 901 TTGACGCCCATCCACCGCAAAACAGGCGGTATGGTGTGACATTACCCACTTCAGTACTG 960
 DB 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 QY 961 TTTATTGCCGGACACGATACTAATCTGGCAAAATCTGGCGCGGCGACTGGAGCTCAACTGG 1020
 DB 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 QY 1021 ACGTTCCTCCGGTCCAGCGGATAACACGCCCGAGGTGGTCAACTGGTGTGTTGAACGCTGG 1080
 DB 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 QY 1081 CGTCGGCTAAGCGATAACACCCAGTGGATTACAGTTTCGCTGCTTCCAGACTTTACAG 1140
 DB 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 QY 1141 CAGATGGCTGATAAAACGCCGCTGTCTATTAATACGCCGCCGCGAGAGGTGAACACTGACC 1200
 DB 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 QY 1201 CTGCGAGGATGCAAGAGCGAAATCGCAGGCGATGTTCTGTTGGCAGGTTTTCACGCAA 1260
 DB 401 LeuAlaGlyCysGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 ATCGTCAATCAAGCAGCATACCGCGGTGCAGTTTG 1296
 DB 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

RESULT 12
 AAB36258
 ID AAB36258 standard; protein; 432 AA.
 XX AAB36258;
 AC AAB36258;
 XX 12-SEP-2003 (revised)
 DT 20-FEB-2001 (first entry)
 XX
 DE R15/APPA plasmid translated sequence.
 XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KW environmental pollution; pig.
 OS Rattus sp.
 OS Escherichia coli.
 OS Chimeric.
 XX WO200064247-A1.
 PN 02-NOV-2000.
 PD
 XX 20-APR-2000; 2000WO-CA000430.
 PF
 XX 23-APR-1999; 99US-0130508P.
 PR (UYGU-) UNIV GUELPH.
 PA
 XX Forsberg CW, Golovan S, Phillips JP;
 PI WPI; 2000-687245/67.
 XX DR N-PSDB; AAC68295.
 DR
 XX Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein.
 PS Disclosure; Fig 18; 152pp; English.
 XX
 XX The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which express phytase in the salivary

CC The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which express phytase in the salivary
 CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
 CC on 12-SEP-2003 to standardise OS field)

XX
 SQ Sequence 432 AA;

Alignment Scores:
 Pred. No.: 8,11e-223 Length: 432
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 95.15% Indels: 0
 DB: 3 Gaps: 0

US-10-034-985-1 (1-1323) x AAB36263 (1-432)

QY 1 ATGAAAGCGATCTTAATCCCAATTTTATCTCTTCTGATTCGTTAAACCCGCAATCTGCA 60
 Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 QY 61 TTGCTCAGAGTGAAGCTGAAGCTGGAAGTGTGTGATTTGTCACTGCTCATGGT 120
 Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
 QY 121 GTGCTGCTCAACACGAGCCAGCACTGATCAGGATGTACCCAGACGATGGCCA 180
 Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 QY 181 ACCTGGCGGTAAACATGGGTGTGGCTGACACCGGNGGTGGTGAAGTAACTCGCTATCTC 240
 Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyLeu 80
 QY 241 GGAATACCAACGCCAGCTCTGCTAGCCAGCGATTTGTCGGGAAAAAGGGCTGCCCG 300
 Db 81 GlyHisTyrglnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 QY 301 CAGCTGTGTCAGTTCGGATTATTGCTGATGTCGAGCGGTACCCGTAACACGAGCGAA 360
 Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 QY 361 GCCTTCCGCGCGGCTGGCACCTGACTGTGCAATACCGTACATACCCAGGAGATACG 420
 Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 QY 421 TCAGTCCCGATCCGTTATTATTCCTCTTAAACCTGGCGGTTTCCCACTGGATACGCG 480
 Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 QY 481 AAGTGTACTCAGCGATCTCCTCAGCAGGCGAGGGTCAATTGCTGACTTTTACCGGCAT 540
 Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 QY 541 CGGCAACCGCGGTTCCGCAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 600
 Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValIleuAsnPheProGlnSerAsnLeuCys 200
 QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAAACGAGGATTAACCATCGGAATC 660
 Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 QY 661 AAGTGTAGCGCCGATGCTCATTTAAACGGTTCGGTGAAGCTTCGATCAATGCTGACG 720
 Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 QY 721 GAGATATTTCTCTGCAACAGCACAGGGAATCCGAGCGGGGTGGGGAAGGATCACCC 780
 Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 QY 781 GATTCACACCGAGTGGAAACACTTGTCTAAGTTTGTGATTAACGCGCAATTTTATTTGCTACAA 840

Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrlLeuGln 280
 QY 841 CGCACCCAGAGGTTCCCGCCAGCCCGCCACCCCGTATTGATTTGATTCATGGCAGCG 900
 Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
 QY 901 TTGACGCCCATCCACCGCAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTACTG 960
 Db 301 LeuThrProHisProGlnLysGlnAlaTyrlGlyValThrLeuProThrSerValLeu 320
 QY 961 TTTATTGCCGACACGATCTAATCTGGCAAAATCTCGGCGCCACTGGAGCTCAACTGG 1020
 Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 QY 1021 ACCTTTCGCGGTGAGCCGATACACCGCCGCGAGGTGGTGAAGTGGTGTGAACTGGT 1080
 Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 QY 1081 CGTCGCTAAGCATAAACAGCCAGTTCAGGTTTCGCTGCTCTTCAGACTTTACAG 1140
 Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 QY 1141 CAGATCGCTGATAAAACGCGCTGCTCATTAATACGCCGCCGAGAGGTGAACCTGACC 1200
 Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 QY 1201 CTGCGCAGGATGTGAAGAGCGAAATGCGCAGGCGATGTTTCGTTGGCAGGTTTTACGCAA 1260
 Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 ATCGTGAATGAAGCAGCATACCGCGCTGCAGTTTG 1296
 Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

RESULT 14
 AAB36260
 ID AAB36260 standard; protein; 432 AA.
 XX
 AC AAB36260;
 XX
 DT 12-SEP-2003 (revised)
 DT 20-FEB-2001 (first entry)
 XX
 DE R15/APPA plasmid translated sequence.
 XX
 DE Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KW environmental pollution; pig.
 KW
 OS Rattus sp.
 OS Escherichia coli.
 OS Chimeric.
 XX
 DN W0200064247-A1.
 XX
 PD 02-NOV-2000.
 XX
 PF 20-APR-2000; 2000WO-CA000430.
 XX
 PR 23-APR-1999; 99US-0130508P.
 XX
 PA (UYGU-) UNIV GUELPH.
 XX
 PI Forsberg CW, Golovan S, Phillips JP;
 XX
 DR WPI; 2000-687245/67.
 DR N-PSDB; AAC68297.
 XX
 XX Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein.
 XX

PS Disclosure; Fig 20; 152pp; English.

XX The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which express phytase in the salivary
 CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
 CC on 12-SEP-2003 to standardise OS field)

XX SQ Sequence 432 AA;

Alignment Scores:

Pred. No.: 8.11e-223 Length: 432
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservativity: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 95.15% Indels: 0
 DB: 3 Gaps: 0

US-10-034-985-1 (1-1323) x AAB36260 (1-432)

QY 1 ATGAAGCGATCTTAATCCATTTTATCTCTTGTGATTCGGTTAAACCCCGCATCTGCA 60
 Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 QY 61 TTCGCTCAGAGTGAGCGGAGCTGAGCTGGAAGTGTGGTATTGTCAGTCTCGTCATGGT 120
 Db 21 PheAlaGlnSerGluProGluLeuLeuLeuGluSerValValIleValSerArgHisGly 40
 QY 121 GTGCTGTCTCCAAAGGCCAGCCCACTGATCAGGATGTCAACCCAGACGGATGGCCA 180
 Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 QY 181 ACTGGCCGTAAACTGGGTTGGCTGACACCGCGGGTGGTGAAGTAAATCGCTATCTC 240
 Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
 QY 241 GGACATTACCAAGCCAGCGCTCTGGTAGCCGCGGATGTCGGCAAAAGGGCTGCCCG 300
 Db 81 GlyHisTyrglnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 QY 301 CAGTCTGGTCAGTTCGGATTATGCTGATGTGCGAGCGGTACCCGTAAACAGCGCAA 360
 Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 QY 361 GCCTTCGCCCGCGGCTGGCACTGACTGTGCAATACCGTACATACCCAGGAGATACG 420
 Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 QY 421 TCAGTTCGCCGATCCGTTATTATTAATCTCTAAACCTGGCGTTTGGCAACTGGATAACGCG 480
 Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 QY 481 AAGTGTACTGACGGATCTTCAGAGGCGAGGAGGTCAATTCCTGACTTTACCGGCGAT 540
 Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 QY 541 CGGCAACCGCGTTTCGGCACTGGAAACGGGTCTTAATTTCCGCAATCAACTGTGTC 600
 Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 QY 601 CTTAAACGTCGAGAAACAGGACGAAAGCTGTTCATTAAACGAGGATTAACCATCGGAATC 660
 Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 QY 661 AAGTGTAGCCCGACAAATGCTCATTAACCGGTGGGTAAAGCTTCGATCAATGCTGACG 720
 Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 QY 721 GACATATTTCTCTGCAACAGCACAGGGAATCCGGAGCCGGGTGGGGAAGGATCAC 780
 Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetCysProGluProGlyTrpGlyArgIleThr 260

QY 781 GATTACACAGTGGAAACCTTGTCTAAGTTTGCATAAAGCGCAATTTATTGCTACAA 840
 Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrlLeuLeuGln 280
 QY 841 CGCACGCCAGAGGTTCGCCGACGCCGCCACCCGTTATTGGAATTGATTCATGGCAGCG 900
 Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
 QY 901 TTGACGCCCATCCACGCCAAAACAGGGGTATGGTGTGACATTACCCACTTCAGTACTG 960
 Db 301 LeuThrProHisProGlnLysGlnAlaTyrglyValThrLeuProThrSerValLeu 320
 QY 961 TTTATTGCCGGACACGATACTATCTGGCAAAATCTCGGGCGGCACCTGGAGCTCAACTGG 1020
 Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 QY 1021 ACGTTCCCGGTGAGCCGGATACACGCCGCCAGGTGGTGAACGTGTGTTGAAACGCTGG 1080
 Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 QY 1081 CGTCGCTAAGCGATAACAGCCAGTCGATTTCAGTTTCGCTGCTTCCAGACTTTACAG 1140
 Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 QY 1141 CAGATCGGTGATAAAACGCCGCTGTCTATTAAATACGCCGCCGAGAGGTGAAACTGACC 1200
 Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 QY 1201 CTGGCAGATGTGAAGAGCGAAATGCGCAGGCGCATGTGTTGTTGGCAGGTTTACGCAA 1260
 Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 ATCGTCAATCAGCACGCATACCGCGTGCAGTTTG 1296
 Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

RESULT 15
 AAU77775
 ID AAU77775 standard; protein; 432 AA.
 XX
 AC AAU77775;
 DT 05-JUN-2002 (first entry)
 XX
 DE Phytase associated protein.
 XX
 KW Phytase.
 XX
 OS Unidentified.
 XX
 PN KR99086028-A.
 XX
 PD 15-DEC-1999.
 XX
 PF 25-MAY-1998; 98KR-00018810.
 XX
 PR 25-MAY-1998; 98KR-00018810.
 XX
 PA (WOJ-) WOJIN CO LTD.
 XX
 PI Bae HD, Forceburgh CW, Goloben S, Cheng KJ;
 XX
 DR WPI; 2000-645078/62.
 DR N-PSDB; ABK12514.
 XX
 PT Novel phytase gene, recombinant phytase and usage thereof.
 PS Disclosure; Fig 3; 10pp; Korean.
 XX
 CC The invention relates to a novel phytase gene, a recombinant phytase gene
 CC and their uses. This is the amino acid sequence of the phytase associated
 CC protein described in the invention

XX Sequence 432 AA;
SQ Alignment Scores:
Pred. No.: 8.11e-223 Length: 432
Score: 2235.00 Matches: 430
Percent Similarity: 99.54% Conservativity: 0
Best Local Similarity: 99.54% Mismatches: 2
Query Match: 95.15% Indels: 0
DB: 3 Gaps: 0

US-10-034-985-1 (1-1323) x AAU77775 (1-432)

QY	1	ATGAAGCGATCTTAATCCATTTTATCTTCTGATTCGGTTAACCCTGCAATCTGCA	60
Db	1	MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla	20
QY	61	TTGCTCAGAGTGCAGCGGAGCTGAGCTGAAAGTGTGATGTCAGTCTGCTGCTG	120
Db	21	PheAlaGlnSerGluProGluLeuLeuLysGluSerValValIleValSerArgHisGly	40
QY	121	GTGCTGCTCAACCAAGGCCAGCAACTGATGAGGATGTCACCCAGACGATGSCCA	180
Db	41	ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro	60
QY	181	ACCTGCGGTAAACCTGGTGTGCTGACACCGGNGGTGGTGAGCTAATCGCTATCTC	240
Db	61	ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu	80
QY	241	GGACATTACCAAGCCAGCTGCTGAGCAGCAGGATGCTGCGGAAAGAGGCTGCCCG	300
Db	81	GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro	100
QY	301	CAGTCTGCTCAGGTCGGGATTATGCTGATGTGCACGAGGCTACCCGTAACACAGGCGAA	360
Db	101	GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu	120
QY	361	GCCTTCGCCCGGGTGGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG	420
Db	121	AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
QY	421	TCCAGTCCCGATCCGTTATTTATCTTAACCTTAACACTGGCGTTTCCCACTGGATACGG	480
Db	141	SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla	160
QY	481	AACGTGACTCAGCGATCTCCTCAGCAGGCGAGGGTCAATTGCTGACTTTTACCGGGCAT	540
Db	161	AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis	180
QY	541	CGGCAAAACGGCGTTCCGCAACTGGAAACGGGTGCTTAATTTCCGCAATCAAACTTGTC	600
Db	181	ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys	200
QY	601	CTTAACGTCAGAAACAGGACGAAGCTGTTCATTACGAGGAGCATTTACCATCGGAATC	660
Db	201	LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu	220
QY	661	AAGGTGAGCCCGCAATGCTCAATTAAACCGTGCCTGAGCTCCGATCAATGCTGACG	720
Db	221	LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr	240
QY	721	GAGATATTTCTCTGCAACAGCAGGGAATCCGAGCGCGGGTGGGAGGATCACC	780
Db	241	GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr	260
QY	781	GATTACACAGTGGAAACCTTGCTAAGTTTGCTAATACCGGCAATTTTATTTGCTACAA	840
Db	261	AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln	280
QY	841	CGCACGCCAGAGGTTCGCCGAGCCGCGACCCCGTTATTGGATTGATGTCATGGCAGCG	900
Db	281	ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla	300

Search completed: March 11, 2005, 23:28:07
Job time : 237.5 secs

QY	901	TTGACGCCCCATCCACGGCAAAACAGGCGTATGTTGATTTACCCACTTCAGTACTG	960
Db	301	LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu	320
QY	961	TTTATTGCGGACACGATACTAATCTGGCAAAATCTCGGCGCGCACTGGAGCTCAACTGG	1020
Db	321	PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp	340
QY	1021	ACGCTTCCCGGTTCAGCCGATACACGCGCCAGGTGGTGAACCTGGTGTGTTGAAACGCTGG	1080
Db	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp	360
QY	1081	CGTCGGCTAAGCGATACACGCCAGTGCAGTTTCAGTTTCGCTGCTTCCAGACTTTACAG	1140
Db	361	ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln	380
QY	1141	CAGATCGGTGATAAAACGCGCTGTCTATTAAATACGCGCCGAGAGGTGAAACTGACC	1200
Db	381	GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr	400
QY	1201	CTGCGAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTCTGTTGGCAGGTTTACGCAA	1260
Db	401	LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln	420
QY	1261	ATCGTCAATGAACACGACATACCGCGCTGCAGTTTG	1296
Db	421	IleValAsnGluAlaArgIleProAlaCysSerLeu	432

THIS PAGE BLANK (USPS 100-100)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 11, 2005, 23:17:45 ; Search time 55.5 Seconds
(without alignments)
4587.199 Million cell updates/sec

Title: US-10-034-985-1
Perfect score: 2349
Sequence: 1 atgaagcgatcttaatccc.....atcaccatcaccatcactaa 1323

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p,model -DEV=xlp
-Q=/cgn2/USFTO.spool_p/US10034985/runat_10032005.164240.11959/app_query.fasta_1.1479
-DB=PIR_79 -QMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOFCU=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPEIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10034985@cgn 1 1 71 @runat 10032005.164240.11959 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2235	95.1	432	2 B36733	acid phosphatase (
2	2210	94.1	434	2 H90770	phosphoanhydride p
3	2210	94.1	444	2 D85631	hypothetical prote
4	943.5	40.2	441	2 AC0201	acid phosphatase (
5	545	23.2	413	2 F90773	periplasmic glucos
6	545	23.2	413	2 D85636	periplasmic glucos
7	541	23.0	417	2 S25627	glucose-1-phosphat
8	539	22.9	413	2 JV0087	glucose-1-phosphat
9	526	22.4	414	2 E87316	periplasmic phosph
10	524	22.3	413	2 AG0632	glucose-1-phosphat
11	139.5	5.9	423	1 S06167	acid phosphatase (
12	133	5.7	416	2 T16058	hypothetical prote
13	131	5.6	423	2 A33395	acid phosphatase (
14	122.5	5.2	421	2 S14742	acid phosphatase (

15	117	5.0	386	1 JH0610	acid phosphatase (
16	116.5	5.0	450	2 A54429	paired box transcr
17	113.5	4.8	344	2 B89130	protein F52E1.8 [i
18	111	4.7	2554	1 TVFF7L	kinase-related pro
19	110	4.7	459	2 S52250	paired box transcr
20	109.5	4.7	1662	2 T18540	mofA protein precu
21	108.5	4.6	10223	2 T30225	polyketide synthas
22	107.5	4.6	457	2 A56925	paired box transcr
23	107	4.6	1048	1 BVECS	exonuclease (EC 3.
24	106.5	4.5	441	2 H70632	hypothetical prote
25	104.5	4.4	2535	2 AC0304	probable hemolysin
26	102	4.3	369	2 JQ2278	hydroxymethylbilan
27	102	4.3	638	2 A29440	signal recognition
28	102	4.3	1047	2 C85535	ATP-dependent dsDN
29	102	4.3	1047	2 G90684	ATP-dependent dsDN
30	101.5	4.3	397	2 C81716	hypothetical prote
31	101.5	4.3	479	1 JN0890	acid phosphatase (
32	101.5	4.3	524	2 H75530	probable acid-CoA
33	101.5	4.3	769	2 B87681	tyrosine kinase Di
34	101.5	4.3	4135	2 T42629	tenascin-X - bovin
35	101	4.3	360	2 JE0116	zinc-finger protei
36	101	4.3	26926	1 I38344	titin, cardiac mus
37	100.5	4.3	438	2 S64682	acid phosphatase (
38	100.5	4.3	478	2 S58315	major plasmid tran
39	100.5	4.3	1214	2 AG2897	conserved hypotet
40	100.5	4.3	1387	2 A97673	probable periplasm
41	100.5	4.3	6260	2 T30228	polyketide synthas
42	100	4.3	328	2 A83572	pyridoxal phosphat
43	100	4.3	1374	1 VCB17	major capsid prote
44	100	4.3	3828	2 T13857	trithorax protein
45	99	4.2	636	2 G81983	piilin glycosylatio

ALIGNMENTS

RESULT 1

B36733

N; Acid phosphatase (EC 3.1.3.2) precursor - Escherichia coli (strain K-12)

C; Species: Escherichia coli

C; Date: 19-Apr-1991 #sequence revision 19-Apr-1991 #text change 09-Jul-2004

C; Accession: B36733; S18018; B64839; A26534; S17960; S33278

J; Dassa, J.; Marck, C.; Boquet, P.L.

R; Bacteriol. 172, 5497-5500, 1990

A; Title: The complete nucleotide sequence of the Escherichia coli gene appA reveals eign

A; Reference number: A36733; MUID:90368616; PMID:2168385

A; Accession: B36733

A; Status: preliminary

A; Molecule type: DNA

A; Residues: 1-432 <DAS>

A; Cross-references: UNIPROT:P07102; GB:M58708; NID:G145283; PIDN:AAA72086.1; PID:G145285

R; Greiner, R.; Jany, K.D.

Biol. Chem. Hoppe-Seyler 372, 664-665, 1991

A; Title: Characterization of a phytase from Escherichia coli.

A; Reference number: S18018

A; Accession: S18018

A; Molecule type: protein

A; Residues: 23-33 <GRE>

R; Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

.A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A; Title: The complete genome sequence of Escherichia coli K-12.

A; Reference number: A64720; MUID:97426617; PMID:9278503

A; Accession: B64839

A; Status: nucleic acid sequence not shown; translation not shown

A; Molecule type: DNA

A; Residues: 1-432 <BLAT>

A; Cross-references: GB:AE000200; GB:U00096; NID:G2367111; PIDN:AAC74065.1; PID:G1787215;

A; Experimental source: strain K-12, substrain MG1655

R; Touati, E.; Danchin, A.

Biochimie 69, 215-221, 1987

A; Title: The structure of the promoter and amino terminal region of the pH 2.5 acid phosph

A; Reference number: A26534; MUID:87271766; PMID:3038201

A;Accession: A26534
A;Molecule type: DNA
A;Residues: 1-50, 'NAGCHPRRMANLAG', 65, 'T', 67-74, 'DV', 77-111, 'S', <TOU>
A;Cross-references: GB:X05471; NID:g40925; PIDN:CAA29031.1; PID:g40927
R;Passa, J.; Fsihi, H.; Marck, C.; Dion, M.; Kieffer-Bontemps, M.; Boquet, P.L.
Mol. Gen. Genet. 229, 341-352, 1991
A;Title: A new oxygen-regulated operon in *Escherichia coli* comprises the genes for a putative phosphatase and a phosphatidylesterase
A;Reference number: S17958; MUID:92049231; PMID:1658595
A;Accession: S17960
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-17 <DA2>
A;Cross-references: GB:S63811; NID:g238656; PIDN:AAB20286.1; PID:g238659
R;Greiner, R.; Konietzny, U.; Jany, K.D.
Arch. Biochem. Biophys. 303, 107-113, 1993
A;Title: Purification and characterization of two phosphatases from *Escherichia coli*.
A;Reference number: S33278; MUID:93256556; PMID:8387749
A;Accession: S33278
A;Molecule type: protein
A;Residues: 23-31, 'A', 33-34 <GR2>
C;Comment: In addition to CAMP-mediated control, this enzyme is induced when bacterial cells are grown under aerobic conditions.
C;Genetics:
A;Gene: appA
A;Map position: 25 min
C;Keywords: monomer; periplasmic space; phosphohistidine; phosphoprotein; phosphoric monoester hydrolase
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-112/Product: acid phosphatase #status predicted <MAT>
F;38/Active site: Arg #status predicted
F;39/Active site: His (phosphohistidine intermediate) #status predicted

Alignment Scores:

Pred. No.:	8, 87e-175	Length:	432
Score:	2235.00	Matches:	430
Percent Similarity:	99.54%	Conservative:	0
Best Local Similarity:	99.54%	Mismatches:	2
Query Match:	95.15%	Indels:	0
DB:	2	Gaps:	0

US-10-034-985-1 (1-1323) x B36733 (1-432)

QY	1	ATGAAGCGATCTTAATCCATTTTATCTCTGATCCGTTAACCCGCAATCTGCA	60
DB	1	MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla	20
QY	61	TTGCTCAGAGTACGCGGAGCTGAAGCTGAAGTGTGTGATTGTCAGTCCTCATGTT	120
DB	21	PheAlaGlnSerGluProGluLeuLysLeuLysLeuValValIleValSerArgHisGly	40
QY	121	GTGCGTCTCAACCAAGGCGCAGCACTGATGAGGATGTACCCGACGCGATGCGCA	180
DB	41	ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTyrPro	60
QY	181	ACCTGCGCGTAAACTGGGTGGCTGACACCGCGGGTGGTGAAGTAATCGCCTATCTC	240
DB	61	ThrTyrProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu	80
QY	241	GGACATTACCAACCGCGCTCTGGTAGCCGAGATGTCACCCGACGCGATGCGCG	300
DB	81	GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro	100
QY	301	CAGTCTGGTTCAGTTCGGATATTGCTGATGTCACGAGGATACCCGTAATAACAGCGAA	360
DB	101	GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu	120
QY	361	GCCTTCGCCCGCGCTGGACCTGATGTGCAATACCGTACATACCCAGGAGATACG	420
DB	121	AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
QY	421	TCAGTCCCGATCCGTTATTTATCTCTAAACTGGCGTTTGGCCACTGGATACGCG	480
DB	141	SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla	160

RESULT 2

phosphoanhydride phosphorylase [imported] - *Escherichia coli* (strain O157:H7, substrain F)
C;Species: *Escherichia coli*
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: H90770
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.;
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H90770
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-434 <HAY>
A;Cross-references: UNIPROT:Q8XC29; GB:BA000007; PIDN:BA834559.1; PID:g13360596; GSPDB:B.1
A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:
A;Gene: Ec81136

Alignment Scores:

Pred. No.: 9.87e-173 Length: 434
Score: 2210.00 Matches: 427
Percent Similarity: 98.39% Conservatives: 0
Best Local Similarity: 98.39% Mismatches: 5
Query Match: 94.08% Indels: 2
DB: 2 Gaps: 1

US-10-034-985-1 (1-1323) x H90770 (1-434)

```
QY 1 ATGAACGATCTTAATCCCATTTTATCTCTTCTGATTCGGTTAAACCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGT-----GAGCCGAGCTGAAGCTGGAAGTGTGGTATTGTCAGTCGT 114
Db 21 PheAlaGlnSerGluProGluProGluLeuLysLeuGluSerValValIleValSerArg 40
QY 115 CATGGTGTGGCTGCTCCAAACCAAGCCACCACTGATGCAGGATGTCAACCCAGAGCGCA 174
Db 41 HisGlyValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAla 60
QY 175 TGCCCAACCTGGCCGGTAAACCTGGTGTGCTGACACCGCGNGGTGTGAGCTAATCGCC 234
Db 61 TrpProAsnTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAla 80
QY 235 TATCTCGGATTAACCAACCCAGCGTCTGGTAGCCGACGAGTGTGGCGGAAAGGGC 294
Db 81 TyrLeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuThrLysLysGly 100
QY 295 TGCCCGCAGCTGTGCTAGGTCGCGATTATTGCTGATGTCCAGCAGCGTACCCGCTAAACAA 354
Db 101 CysProGlnProGlyGlnValAlaIleLeuAlaAspValAspGluArgThrArgLysThr 120
QY 355 GGCGAACCTTGGCCGCGCGGTGGACCTGCTGCAATAACCGTACATACCCAGGCA 414
Db 121 GlyGluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAla 140
QY 415 GATACGTCGATCCGATCCGTTATTTAATCCTCTAAAACTGGCGTTCCCACTGGAT 474
Db 141 AspThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAsp 160
QY 475 AACCGAACGTGACTGACGCGATCCTCAGCAGGCGAGGCTCAATTGCTGACTTTACC 534
Db 161 AsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThr 180
QY 535 GGGCATCGCAACCGCGGTTTCGCGAACTCGAACGGGTGCTTAATTTTCGCAATCAAC 594
Db 181 GlyHisArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsn 200
QY 595 TTGTGCTTAAACGTGAGAAACAGGACGAAAGCTGTTTCATTAACGAGGATACCATCG 654
Db 201 LeuCysLeuAsnArgGlyLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSer 220
QY 655 GAATCAAGGTGACGCGCGCAATGTCTCTTAACCGGTGCGTAAGCCTCGCATCAATG 714
Db 221 GluLeuLysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMet 240
QY 715 CTGACGGAGATATTTCCTCTGCAACAAAGCAGGAAATGCGGAGCCGGGTGGGGAAGG 774
Db 241 LeuThrGluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArg 260
QY 775 ATCAACGATTCACACGATGGAACACCTTGCTAAGTTTGATTAACCGGCAATTTATTG 834
Db 261 IleThrAspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeu 280
QY 835 CTACAACGACGCGCAGAGTTGCGCGAGCGCGCCACCCGCTTATTGGATTTGATCATG 894
Db 281 LeuGlnArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMet 300
```

```
QY 895 GCAGCGTTCAGCGCCCATCCACCGCAAAACAGCGCTATGGTGTGACATTACCACTTCA 954
Db 301 IleAlaLeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSer 320
QY 955 GTACTGTTTATTCGCGACACGATACTAATCTCGCAAACTCTCGCGCGCGACCTGAGCTC 1014
Db 321 ValLeuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeu 340
QY 1015 AACTGACGCTTCCCGGTAGCCGATAAACACCCGCCAGGTGGTGAACCTGGTGTGAA 1074
Db 341 AsnTrpThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGlu 360
QY 1075 CGCTGCGCTCGCTAAGCATACACGACGATTCAGGTTCGGTTCGTTCCAGACT 1134
Db 361 ArgTrpArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThr 380
QY 1135 TTACAGCAGATGGTGATAAACCGCGCTGTCTAATAAATACGCGCGCGAGAGGTGAAA 1194
Db 381 LeuGlnGlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLys 400
QY 1195 CTGACCTGGCAGGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTGTTGGCAGGTTT 1254
Db 401 LeuThrLeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPhe 420
QY 1255 ACCCAATCGTGAATGACGACGATACCGCGCTGCAGTTTG 1296
Db 421 ThrGlnIleValAsnGluAlaArgIleProAlaCysSerLeu 434
```

RESULT 3

D85633
hypothetical protein appA [imported] - Escherichia coli (strain O157:H7, substrain EDL93:
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: D85633
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouasis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85633
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-444 <STO>
A;Cross-references: UNIPROT:Q8XC29; GB:AE005174; MID:g12514245; PIDN:AAG55528.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: appA

Alignment Scores:

Pred. No.: 9.9e-173 Length: 444
Score: 2210.00 Matches: 427
Percent Similarity: 98.39% Conservatives: 0
Best Local Similarity: 98.39% Mismatches: 5
Query Match: 94.08% Indels: 2
DB: 2 Gaps: 1

US-10-034-985-1 (1-1323) x D85633 (1-444)

```
QY 1 ATCAAGCGATCTTAATCCCATTTTATCTCTTCTGATTCGGTTAAACCCGCAATCTGCA 60
Db 11 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 30
QY 61 TTCGCTCAGAGT-----GAGCCGAGCTGAAGCTGGAAGTGTGGTATTGTCAGTCGT 114
Db 31 PheAlaGlnSerGluProGluProGluLeuLysLeuGluSerValValIleValSerArg 50
QY 115 CATGGTGTGGCTGCTCCAAACCAAGGCCACCGCAACTGATGAGGATGTCAACCCAGAGCGCA 174
Db 51 HisGlyValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAla 70
QY 175 TGCCCAACCTGGCCGGTAAACCTGGTGTGCTGACACCGCGNGGTGTGAGCTAATCGCC 234
Db 71 TrpProAsnTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAla 90
```

```
QY 235 TATCTCGACATTACCAACGCCAGCGTCTGTAGCCGACGGATTGCTGGCGAAAAAGGCG 294
Db 91 TyrLeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuThrLysGly 110
QY 295 TGCCCGCAGTCTGCTAGGTCGCGATTATTCTGATCTGCACGAGCGTACCCGTAACAA 354
Db 111 CysProGlnProGlyGlnValAlaIleAlaAspValaAspGluArgThrArgLysThr 130
QY 355 GCGGAAGCCTTCGCGCGCGGCTGCGACCTGCTGCTGCAATAACCGTACATACCAGGCA 414
Db 131 GlyGluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAla 150
QY 415 GATACGTCAGTCCGATCCGTTATTATTAATCTCTAAAACTGGCGTTTCGCAACTGGAT 474
Db 151 AspThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAsp 170
QY 475 AACCGAAGCTGACTGACGCGATCCTCAGCAGGCGAGGTCATTTGCTGACTTTACC 534
Db 171 AsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThr 190
QY 535 GGGCATCGGCAACCGGCTTCGGAACCTGGAACGGTCTTAATTTTCGCAATCAAC 594
Db 191 GlyHisArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsn 210
QY 595 TTGTGCTTAAACGCTGAGAAACAGCAGCAAGCTGTTCAATTAACGACGATACCATCG 654
Db 211 LeuCysLeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSer 230
QY 655 GAATCAAGGTGACGCGCAATGTCTCAATTAACCGGTGCGTAAACCTCGCATCAATG 714
Db 231 GluLeuLysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMet 250
QY 715 CTGACCGGATATTTCCTGCAACAGCAGGGAATCGGAGCGCGGTCGGGAGG 774
Db 251 LeuThrGluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrpGlyArg 270
QY 775 ATCAACCATTCACACCACTGGAACACCTGTGTAAGTTTGATACCGCGCAATTTATTG 834
Db 271 IleThrAspSerHisGlnTyrAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeu 290
QY 835 CTACAACGACGCGCAGAGTTGCGGACGCGCGCCACCCGTTATTGGATTGATCATG 894
Db 291 LeuGlnArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMet 310
QY 895 GCAGGTTGACGCGCCATCCACCGCAAAACAGCGCATGTGTGACATTACCCACTCA 954
Db 311 IleAlaLeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSer 330
QY 955 GTACTGTTTATTGCGGACACGATACTAATCTGGCAAAATCTCGCGCGCGCACTGGAGCTC 1014
Db 331 ValLeuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeu 350
QY 1015 AACTGAGCGTTCCCGTACGCGGATAACACGCGCGCGAGTGTGTGAACTGGTTTGA 1074
Db 351 AsnTrpThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGlu 370
QY 1075 CGCTGGCGTGGCTAAGCGATAACAGCAGTGGATTTCAGGTTTCGCTGCTCCAGACT 1134
Db 371 ArgTrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThr 390
QY 1135 TTACAGCAGATGCTGTATAAACCCGCTGTCTATTAAATACGCCCGCGAGAGGTGAAA 1194
Db 391 LeuGlnGlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLys 410
QY 1195 CTGACCCCTGGCAGATGTGAAGACGAAATCGCAGGCGCATGTGTTCTGGCGAGTTT 1254
Db 411 LeuThrLeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPhe 430
QY 1255 ACGCAATCGTGAATGAGCAGCATACCGCGCTGAGTTTG 1296
Db 431 ThrGlnIleValAsnGluAlaArgIleProAlaCysSerLeu 444
```

RESULT 4

AC0201 acid phosphatase (EC 3.1.3.2) [imported] - Yersinia pestis (strain CO92)

C;Species: Yersinia pestis

C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C;Accession: AC0201

R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; F

il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, I

Nature 413, 523-527, 2001

A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A;Reference number: AB0001; MUID:21470413; PMID:11586360

A;Accession: AC0201

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-441 <R>

A;Cross-references: UNIPROT:Q8ZFP6; GB:AL590842; PIDN:CAC90470.1; PID:gl5979685; GSPDB:G

C;Genetics:

C;Gene: YPO1648

C;Keywords: phosphoric monoester hydrolase

Alignment Scores:

Pred. No.: 4,5e-69 Length: 441

Score: 943.50 Matches: 199

Percent Similarity: 60.78% Conservative: 66

Best Local Similarity: 45.64% Mismatches: 156

Query Match: 40.17% Indels: 15

DB: 2 Gaps: 7

US-10-034-985-1 (1-1323) x AC0201 (1-441)

QY 13 TTAATCCCAATTTTATCTCTCTGATTTCGTTAAACCCGCAATCTGCAATTCGTCAGAGT 72

Db 13 LeuValLeuMetLeuSerGlyLeuAlaAlaIleThrAlaProValAla 29

QY 73 GAGCCGAG---CTGAGCTGGAAGTGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCT 129

Db 30 GluProSerGlyTyrThrLeuGluArgValIleLeuSerArgHisGlyValArgSer 49

QY 130 CCAACCAAGCCACCACTGATGAGGATGTCACCCAGCATGCGCCCACTGCGCCG 189

Db 50 ProThrLysGlnThrGlnLeuMetAsnAspValThrProAspLysTrpProGlnTrpPro 69

QY 190 GTAAACTGGGTTGGCTGACACCGCGNGTGGTGGAGTAATCCCTATCTCGACATTC 249

Db 70 ValLysAlaGlyTyrLeuThrProArgGlyAlaGluLeuValThrLeuMetGlyGlyPhe 89

QY 250 CAAGCCAGCGTCTGTAGCCGAGGATTCCTGCGGAAAAAGGCTCCCGCAGTCTGCT 309

Db 90 TyrGlyAspTyrPheArgSerLeuGlyLeuAla---AlaGlyCysProAlaGluGly 108

QY 310 CAGTCCGATTTATTGCTGATGTCGAGCGTACCCGTAAAAACAGCGAGGCTTCGCC 369

Db 109 GlyValTyrAlaGlnAlaAspIleAspGlnArgThrArgLeuThrGlyGlnAlaPheLeu 128

QY 370 GCCGGTGGCACCTGCTGCAATAACCGTACATACCCAGGAGATACGTCAGTCCC 429

Db 129 AspGlyValAlaProGlyCysGlyLeuThrValHisAsnGlnAlaAspLeuLysThr 148

QY 430 GATCCGTTATTTATCTCTAAAAACCTGGGTTTGGCACTGGATTAACGCGACGTGACT 489

Db 149 AspProLeuPheHisProValGluAlaGlyValCysLysLeuAspAlaAlaGlnThrAsp 169

QY 490 GAGCCGATCTCAGCAGGAGGTCATCTGCTGACCTTTACCGGCGCATCGCAACG 549

Db 169 LysAlaIleGluGlnLeuGlyGlyProLeuAspThrValSerGlnArgTyrAlaLys 188

QY 550 GCGTTTCGCACTGGAACCGGTCGTTAATTTCCGCAATCAAACTTGCTGCTTAAACGT 609

Db 189 ProPheAlaGlnMetGlyAspValLeuAsnPheAlaAsaProTyrCysLysSerLeu 208

QY 610 GAGAAACAGGACGAAAGCTGTTTCATTAAACCGAGCATTAACGAGGTCAGGTCAGC 669


```

Db      332 GlnAsnGluArgThrProIleGlyGlyLeuValPheGlnArgThrArgAspSerLys 351
QY      1093 GATAACAGCCAGTGGATTTCAGTTTCCTTCAGACTTTTACAGCAGATCGGTGAT 1152
Db      352 AlaAsnArgAspLeuMetLysIleGluValTyrGlnSerAlaGluGlnLeuArgAsn 371
QY      1153 AAAACCGCGTGTCTATTAAATACCGCCCGGAGAGGTGAACCTGACCTGGCAGGATGT 1212
Db      372 AlaAspAlaLeuThrLeuGlnAlaProAlaGlnArgValThrLeuGluLeuSerGlyCys 391
QY      1213 GAAGAGCGAAATGCGCAGGCGCATGTTCGTGTGCAGCTTTTACGCCAAATCGTGAATGAA 1272
Db      392 ---ProIleAspAlaAspGlyPheCysProMetAspLysPheAspSerValLeuAsnGlu 410
QY      1273 GCA 1275
Db      411 Ala 411

RESULT 6
B85636
periplasmic glucose-1-phosphatase [imported] - Escherichia coli (strain O157:H7, substra
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: B85636
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: B85636
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-413 <STO>
A;Cross-references: UNIPROT:Q8XBZ6; GB:AE005174; NID:gl2514273; PIDN:AAG55550.1; GSPDB:G
A;Experimental source: strain O157:H7, substrate EDL933
C;Genetics:
A;Gene: agp

Alignment Scores:
Pred. No.: 1.84e-36 Length: 413
Score: 545.00 Matches: 135
Percent Similarity: 50.12% Conservative: 76
Best Local Similarity: 32.07% Mismatches: 180
Query Match: 23.20% Indels: 30
DB: 2 Gaps: 10

US-10-034-985-1 (1-1323) x B85636 (1-413)
QY      37 ATTCGGTTAAACCCCGCAATCTGCATTCGCTCAGAGTCAGCGGAG---CTGAAGCTGGAA 93
Db      13 ValValLeuLeuAlaSerAsnAlaGlnAlaGlnThrValProGluGlyTyrGlnLeuGln 32
QY      94 AGTGTGTGATGTCAGTCGTATGTCGTGTGTCCTCCA---ACCAAGGCCACGCAACTG 150
Db      33 GlnValLeuMetSerArgHisAsnLeuArgAlaProLeuAlaAsnAsnGlySerVal 52
QY      151 ATGCAGAGATGTCACCCAGCAGCATGCGCAACCTGCGCGGTAAACCTGGTGTGCTGACA 210
Db      53 LeuGlnSerThrProAsnLysTrpProGluTrpAspValProGlyGlyGlnLeuThr 72
QY      211 CCGCGNGTGTGAGTAAATCGCTATCTCGACATACCAACGCCAGCGTCTGGTAGCC 270
Db      73 ThrLysGlyValLeuGluValTyrMetGlyHisTyrMetArgGluTrpLeuAlaGlu 92
QY      271 GACGGATTGTCGGCAAAAAGGGCTGCCCGCAGTCTCGTTCAGTTCGCGATTTATGCTGAT 330
Db      93 GlnGlyMetValLysSerGlyGluCysProProAspThrValTyrAlaTyrAlaAsn 112
QY      331 GTCGACAGGCTACCCGTAACACAGCGAAGCCCTTCGCCCGCGCTGGCACCTGACTGT 390
Db      113 SerLeuGlnArgThrValAlaThrAlaGlnPheIleThrGlyAlaPheProGlyCys 132
QY      391 GCAATACCGGTACATACCCAGGCAGATACGTCCAGTCCCGATCCGATTCGCTTATTATTCCTCTA 450

```

RESULT 7

S25627

glucose-1-phosphatase (BC 3.1.3.10) precursor - Providencia rettgeri

C;Species: Providencia rettgeri

C;Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004

A;Accession: S25627

R;Riccio, M.L.; Chiesurin, A.; Lombardi, G.; Satta, G.

submitted to the EMBL Data Library, September 1992

A;Reference number: S25627

A;Accession: S25627

```

Db      133 AspIleProValHisIleGlnGluYMetGlyThrMetAspProThrPheAsnProVal 152
QY      451 AAAACTGGCGCTTTGCCAACTGAGTACGGAACGTGACTACGCGATCCTCAGCAGGGCA 510
Db      153 IleThrAsp-----AspSerAlaAlaPheSerGluGlnAlaValAlaAlaMet 168
QY      511 GGAGGTCAATTCTGACTTACCGGGCATCGGCAACGGCGTTTCGCGAATCGGAACGG 570
Db      169 GluLysGluLeuSerLysLeu-----GlnLeuThrAspSerTyrGlnLeuLeuGluLys 186
QY      571 GTCTTAATTTCCGCAATCAAACTTGTGCTTAAACGTGAGAAACAGCAGCAAGCTGT 630
Db      187 IleValAsnTyrLysAspSerProAlaCys-----LysGluLysGlnGlnCys 202
QY      631 TCATTAACGAGCAGCATTACCATCGGAACCTCAAGGTGAGCCGCCACAAATGTCTCATTAAC 690
Db      203 SerLeuValAspGlyLysAsnThrPheSerAlaLysTyrGlnGlnGluProGlyValSer 222
QY      691 GGTGCGGTAACTCGCATCAATGCTGACGAGATATTTCTCTGCAACAAAGCACAGGGA 750
Db      223 GlyProLeuLysValGlyAsnSerLeuValAspAlaPheThrLeuGlnTyrTyrGluGly 242
QY      751 ATGCCG-----GAGCGGGGTGGGAAGGATCACCGATTTCACACGAGTGAACACCTTG 804
Db      243 PheProMetAspGlnValAlaTyrGlyGluLysSerAspGlnGlnTyrLysValLeu 262
QY      805 CTAAGTTTGCATTAACCGGCANNTTATTTGCTACAACGCGCAGCAGGTTGCCGCGAGC 864
Db      263 SerLysLeuLysAsnGlyTyrGlnAspSerLysPheThrSerProGluValAlaAlaArgAsn 282
QY      865 CGCGCCACCCCGTTATTGGATTGATCATCGGACGCTTGACGCCCATCCACCGCAAAA 924
Db      283 ValAlaLysProLeuValSerTyrIleAspLysAlaLeu----- 295
QY      925 CAGGCGTATGCTGTGACATTACCCACTTCA-----GTACTGTTTATTTCGCGGACAC 975
Db      296 -----ValThrAspArgThrSerAlaProLysIleThrValLeuValGlyHis 311
QY      976 GATACTAATCTGCAATCTCGCGGGGACACTGAGGCTCAAC---TGGACGCTTCCCGGT 1032
Db      312 AspSerAsnIleAlaSerLeuThrAlaLeuAspPheLysProTyrGlnLeuHisAsp 331
QY      1033 CAGCGGATAACACACGCGCGCAGCTGTTGAACTGCTGTTTCAACGCTGCGCTGCGCTAAGC 1092
Db      332 GlnAsnGluArgThrProIleGlyLysIleValPheGlnArgTrpArgSerLys 351
QY      1093 GATAACAGCCAGTGGATTTCAGGTTTCGCTGCTTCAGACTTTTACAGCAGATCGGTGAT 1152
Db      352 AlaAsnArgAspLeuMetLysIleGluTyrValTyrGlnSerAlaGluGlnLeuArgAsn 371
QY      1153 AAAACCGCGTGTCTAATAATACGCGCGCGGAGAGGTGAACCTGACCTGGCAGGATGT 1212
Db      372 AlaAspAlaLeuThrLeuGlnAlaProAlaGlnArgValThrLeuGluLeuSerGlyCys 391
QY      1213 GAAGAGCGAATCGCAGGCGCATGTGTTGTCGAGGTTTACGCAAAATCGTGAATGAA 1272
Db      392 ---ProIleAspAlaAspGlyPheCysProMetAspLysPheAspSerValLeuAsnGlu 410
QY      1273 GCA 1275
Db      411 Ala 411

```


A:Molecule type: DNA
A:Residues: 1-417 <RC>
A:Cross-references: UNIPROT:Q52309; EMBL:X68201; NID:945771; PIDN:CAA48288.1; PID:945772
A:Experimental source: strain PV7
C:Genetics:
A:Gene: agp
C:Keywords: periplasmic space; phosphohistidine; phosphoprotein; phosphoric monoester by
F;1-29/Domain: signal sequence #status predicted <Sig>
F;30-417/Product: glucose-1-phosphatase #status predicted <MAT>
F;42/Active site: His (phosphohistidine intermediate) #status predicted
F;314/Active site: His #status predicted

Alignment Scores:
Pred. No.: 3.92e-36 Length: 417
Score: 541.00 Matches: 129
Percent Similarity: 50.35% Conservative: 84
Best Local Similarity: 30.50% Mismatch: 182
Query Match: 23.03% Indels: 28
Gaps: 9

US-10-034-985-1 (1-1323) x S25627 (1-417)

QY 28 TCCTCTCGATTCGGTAAACCCCGCAATCTGCTCAGAGTGAGCGGAGCTGAAG 87
DB 13 AlaleuphealaproiliealaprothrmetalaSerThrAspAenGlnAlaAspMetVal 32
QY 88 CTGGAAGTGTGTGATGTGTCAGTCGTGTCGTGTCGTCTCCAAACGAGCGGAGCA 147
DB 33 LeuAspGlnValLeuValLeuSerArgHisAenLeuArgThrProilieValAenThrGly 52
QY 148 CTGATGAGATGATCACCAGCAGCATGGCAACCTGGCGGCTAAACTGGTGGTGGCTG 207
DB 53 IleLeuThrGluValThrAspLysLysLysLysLysLysLysLysLysLysLysLysLys 72
QY 208 ACACCGGNGTGTGTGAGTAAATCGCTATCTCGACATTACCAACCCAGCGTCTGTA 267
DB 73 ThrThrGlnGlyAlaLeuGluValThrMetGlyHisThrPheArgGluThrPheAsp 92
QY 268 GCCGACGGATGTGCGGAAAGGGCTGCCGAGTCTGTGTGAGTCTGTCAGTCTGTCG 324
DB 93 GlnAenLysLeuLeuAlaAspGluLeuCysProThrSerAenGluAspIleThrLeuTyr 112
QY 325 GCTGATGTCAGAGGATCCGCTAAACAGCGGAGCGTTCGCGCGGCGGTGGACCT 384
DB 113 ThrAenSerLeuGlnArgThrIleAlaThrAlaGlnPhePheAlaAlaGlyAlaPhePro 132
QY 385 GACTGTGCAATACCGTACATACCCAGCAGATAGTCCAGTCCCGATCCGTTATTTAAT 444
DB 133 GlyCysLysValAenIleHisGlnProGluIleGlyLysMetAspProvalPheAen 152
QY 445 CCTCTAAAAACTGGCGTGTGCAACTGATAACCGCAACGTGACTGACGCGATCCTCAGC 504
DB 153 ProfileleThr-----AsnGlySerProGluPheLysGln 164
QY 505 AGGCGAGGAGGTCAATGTGATTTACCGGG-----CATCGGCAACGGCGTTTCGC 558
DB 165 LysAlaLeuAlaAlaMetAspTyrLeuLysGlyLeuSerLeuLysAlaGlyTyrGlu 184
QY 559 GAAGTGAACGGGTGCTTAATTTCCGCAATCAAACTTGTCCTTAAACGTGAGAACAG 618
DB 185 GluLeuAspThrValLeuAenIleLysAspSerGlnLysCys-----LysThr 200
QY 619 GACGAAAGCTGTTTCATTAACGAGCGGATTCATTCGGAACCTCAAGGTGAGCGCGACAAT 678
DB 201 AspLysLeuCysAenLeuAspSerGlnLysAenSerPheIleIleGluAlaAspLysGlu 220
QY 679 GTCTCATTAACGGTGGCGGTGAGCTGCGATCAATGCTGAGCGAGATATTTCTCTGCAA 738
DB 221 ProGlyValSerGlyProLeuLysIleAlaAenSerAlaValAspAlaIleAspLeuGln 240
QY 739 CAAGCAGGAGATGCG-----GACCGGGGTGGGAGGATCACCAGTTCACACCAG 792
DB 241 TyrThrGluGlyPheProAlaAspGlnValAlaTrpGlyLeuValAspThrProGluLys 260

QY 793 TGAACACACCTTGCTAAGTTTGATACAGCGCAATTTTATTGCTACACCGCCAGAG 852
DB 261 TrpLysLeuAenThrLeuLysAenAlaTyrGlnGluThrLeuPheThrProLysIle 280
QY 853 GTTGCCCGCAGCGCGCCACCCCGTTATTGATTTGATTGATCATGCGAGCGTTGAGCCCAT 912
DB 281 IleAlaLysAenValAlaHisProIleLysAenThrIleAspLysGlyPheValSerVal 300
QY 913 CCACCGCAAAACAGCGCGTATGTCACATTACCCACCTTCAGTACTGTTTATTGCGGGA 972
DB 301 AspLysGlyGluThrAla-----LysPheIlePheLeuValGly 313
QY 973 CACGATACCTAATCTGCAAAATCTCGGCGGCGACTGAGAGCTCAAC---TGACGCTTCCC 1029
DB 314 HisAspSerAenIleAlaSerLeuMetSerAlaMetAspPheLysProTyrGlnLeuAla 333
QY 1030 GGTGACCGGATAACAGCGCGAGTGTGTGAACCTGTTGTAAGCTGCGTGGCGTGGCTA 1089
DB 334 GlnGlnTyrGluHisThrProIleGlyLysLeuValPheGlnArgThrAspLys 353
QY 1090 AGCGATAACAGCCAGTGGATTTCAGTTTCGCTGCTTCAGAGCTTTACAGCAGATCGT 1149
DB 354 GlnThrLysLysAspPheMetLysValGluTyrValTyrGlnThrAlaAspGlnLeuArg 373
QY 1150 GATAAAGCGCGTGTCTAATTAATAGCCCGCGGAGAGGTGAACCTGACCTGGCAGGA 1209
DB 374 AspAenAlaTyrLeuSerLeuGluThrProLysHisValThrLeuGluLeuLysAsp 393
QY 1210 TGTGAA--GAGCGAATGCGAGGAGTGTGTTGTCGAGTGTTCGCAAACTCGTG 1266
DB 394 CysProValAspLysAen-----GlyTyrCysSerTrpGluAspPheGlnLysValMet 411
QY 1267 AATGAAGCA 1275
DB 412 AlaLysAla 414

RESULT 8
JV0087
glucose-1-phosphatase (EC 3.1.3.10) precursor - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C:Accession: JV0087; H64841
R:Pradel, E.; Marck, C.; Boquet, P.L.
J. Bacteriol. 172, 802-807, 1990
A:Title: Nucleotide sequence and transcriptional analysis of the Escherichia coli agp
A:Reference number: JV0087; MUID:90130318; PMID:2153660
A:Accession: JV0087
A:Molecule type: DNA
A:Residues: 1-413 <PRA>
A:Cross-references: UNIPROT:P19926; GB:M33807; NID:9145217; PIDN:AAA23426.1; PID:9145218
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
-A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: H64841
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-413 <BLAT>
A:Cross-references: GB:AE000202; GB:U00096; NID:91787233; PIDN:AACT4087.1; PID:91787237;
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: agp
A:Map position: 23 min
C:Function:
A:Description: essential for growth in a high-phosphate medium containing glucose-1-phosph
A:Note: Optimal at low pH
C:Keywords: homodimer; periplasmic space; phosphohistidine; phosphoprotein; phosphoric m
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-413/Product: glucose-1-phosphatase #status predicted <MAT>
F;40/Active site: His (phosphohistidine intermediate) #status predicted
F;311/Active site: His #status predicted

[illegible]

```

Db      81  ArgHisTyrAlaAlaGlnGlyLeuLeuYsProGlyAspCys-----AlaSerVal  97
Qy      316  GCGATTATTGCTGATGTCGACGAGCTACCGTAAACACAGCGAAGCTTCGCGCGCGG  375
Db      98  TyrAlaTrpAlaAsnValThrGlnArgThrIleAlaThrAlaLysAlaTyrArgGluThr  117
Qy      376  CTGCGACCTGACTGTGCAATAACCGTACATACCAGGAGATAGTCCAGTCCGATCCG  435
Db      118  LeuAlaProGlyCysProValThrValAsnThrValGlyGlu---GlyAsnIleAspPro  136
Qy      436  TTATTTAATCTCTAATAACCTGGGTTGCGCACTGATAACGCGAAGCTGACTGACGCG  495
Db      137  MetPheGluProValLysAlaGlyIleValLysAlaAspHisAlaLeuAlaAa-gAlaAla  156
Qy      496  ATCTCTACGAGGCGAGCGGTCAATTGCTGACTTTACCGGCGATCGGCAACGCGCTTT  555
Db      157  ValAlaGlyArgValGlyGlyAspLeuThrAlaTrpSerAlaSerHisAsnGlnGluAla  176
Qy      556  CGGAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTCCTT-----  603
Db      177  GluGlnLeuAspAlaLeuLeuMetGlnCysAspLysGlyProCysProAlaProGly  196
Qy      604  AAGCTGAGAAACGAGGAGAAAGCTGTTCATTACGAGGAGCATTTACCATCGGAACG  663
Db      197  LysArgArgValPheAspAlaLysProGlyPheValAspGly-----GluGluLeu---  213
Qy      664  GTGAGCGCCGACATGCTCATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACGGAG  723
Db      214  -----AlaGlyLeuSerGlyProGluAlaPheAlaSerGlyValThrGlu  228
Qy      724  ATATTTCTCTGCAACAGACAGGGAATCGCGAGCGGGGTGGGA---AGGATCAC  780
Db      229  SerLeuLeuMetAlaTrpAlaAspGlyArgAspPheAlaGlyLeuGlyTyrLysSerLeu  248
Qy      781  GATTACACAGTGAACACCTTGCTAAGTTGATTAACCGGCAATTTTATTATTCTACAA  840
Db      249  AspGluGluAlaLeuThrArgSerPhePheLeuHisGlnAlaGluPheAspLeuArgLeu  268
Qy      841  CGCAGCGCAGAGGTGCGCGAGCGCGCCAGCCCGTTATTGATTGATTCATGGCAGCG  900
Db      269  ArgThrProTyrValAlaAlaGlnThrLeuAlaGlyHisLeuAlaAspArgLeuAlaAlaThr  288
Qy      901  TTGACGCGCCCATCCACGCAAAACAGCGGTATGTTGATGATTTACCCACTTCAGTACTG  960
Db      289  LeuArg-----AspGlyAlaAlaAlaIleGly---ProValAspAlaArgLeuVal  304
Qy      961  TTTATTGCGGACGATACTAATCTCGCAATCTCGCGCGGCGACCTGGAGCTCAACTGG  1020
Db      305  IleIleAlaGlyHisAspGlyThrLeuAlaSerLeuGlyGlyLeuLeuArgMetGluTrp  324
Qy      1021  ACGTCTCCCGGT---CAGCGGATAACACCGCGGTGTTGAACG  1077
Db      325  ThrLeuProGlyTyrGlnProAsnGlnIleGlnProGlyAlaLeuValPheGluArg  344
Qy      1078  TGGCGTCCGCTAAGCGATAACAGCAGTGGATTTCAGTGTTCGCTGCTTCACGACTTTA  1137
Db      345  TrpArgArg---AspAspGlyValArgValValArgValArgPheThrGlyGlnSerLeu  363
Qy      1138  CAGCAGATGGGTGATAAAACCGCGCTGTCAATTAATACGCGCGCGGAGAGGTGAAC  1197
Db      364  SerGlnLeuArgAsnMetThrAlaLeuAspAlaLysThrProProLeuSerAlaProVal  383
Qy      1198  ACCCTGCGAGGATCGAAGAGCGAAATGCGCGGCGCATGTTGCTGTTGGCAGGTTTACG  1257
Db      384  PheValGlnGlyCysGlyThrAlaThrProAlaPheAspCysArgLeuGluAspPheGlu  403
Qy      1258  CAAATCGTGAATGAAGCA  1275
Db      404  ThrValValArgGlyAla  409

```

RESULT 10
AG0632

glucose-1-phosphatase precursor (G1Pase), secreted [imported] - Salmonella enterica subsp.
C;Species: Salmonella enterica subsp. enterica serovar typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AG0632
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Gaora, P.
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
Nature 413, 848-852, 2001
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AG0632
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-413 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD08242.1; PID:gl6502289; GSPDB:GN00176
C;Genetics:
A;Gene: STY1153

Alignment Scores:

Pred. No.:	9,646-35	Length:	413
Score:	524.00	Matches:	135
Percent Similarity:	48.90%	Conservative:	65
Best Local Similarity:	33.01%	Mismatches:	185
Query Match:	22.31%	Indels:	24
DB:	2	Gaps:	10

US-10-034-985-1 (1-1323) x AG0632 (1-413)

```

Qy      64  GTCAGAGTGAAGCGGAG---CTGAAGCTGGAAGTGTGTGATGTCATGTCATGCT  120
Db      22  AlaGlnThrThrProGluGlyTyrGlnLeuGlnValLeuMetSerArgHisAsn  41
Qy      121  GTCGCTGCTCCA---ACCAAGGCCGCACTGATCGAGATGTCACCCAGCAGCATG  177
Db      42  LeuArgAlaProLeuAlaAsnAsnGlyAsnValAlaGlnSerThrProAsnAlaTrp  61
Qy      178  CCAACCTGGCGGTAATAAATGGGTGTGTCACACCGCGNGTGTGAGCTTAATCGCCTAT  237
Db      62  ProAlaTrpAspValProGlyGlyGlnLeuThrThryGlyGlyValLeuGluValTyr  81
Qy      238  CTCGACATTAACAACGCCAGCTCTGTGTAGCCGAGGATGCTGGCGAAAAAGGGTGC  297
Db      82  MetGlyHisTyrThrArgGluTrpLeuValAlaGlnGlyLeuIleProSerGlyGluCys  101
Qy      298  CCGCAGTCTGCTCAGTCCGATTTATGCTGATGTCGACGAGCGTACCCGTAAACAGGC  357
Db      102  ProAlaProAspThrValTyrAlaTyrAlaAsnSerLeuGlnArgThrValAlaThrAla  121
Qy      358  GAAGCCTTCGCCCGCGGTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCAGAT  417
Db      122  GlnPhePheIleThrSerAlaPheProGlyCysAspIleProValHisHisGlnGluLys  141
Qy      418  ACCTCCAGTCCCGATCCGTTATTATCTCTTAAACCTGCGCTTGGCCAACTGGATAAC  477
Db      142  MetGlyThrMetAspProThrPheAsnProValIleThrAspAspSerAlaAlaPheArg  161
Qy      478  GCGAACTGACTCAGCCGATCTCAGCAGGCGAGGGGTCAATTGCTGACTTTACCGGG  537
Db      162  GlnGlnAlaValGlnAla---MetGluLysAlaArgSerGlnLeu-----  175
Qy      538  CATCGCAAAACGGGTTTTCGCAACTGGAACGGGTGCTTAATTTTCGCAATCAAACTTG  597
Db      176  HisLeuAspLysSerTyrLysLeuGlnGlnIleThrHisTyrGlnAspSerProSer  195
Qy      598  TGCCTTAAACGTGAGAAACAGCAGAAAGCTGTTTCAATTAACGAGCATTTACCATCGAA  657
Db      196  Cys-----LysGluLysHisGlnCysSerLeuIleAspAlaLysAspThrPhe  211
Qy      658  CTCAGGTGAGCGCCGACAAATGCTCATTAACCGGTGCGGTAGCCCTCGCATCAATGCTG  717
Db      212  SerAlaAsnTyrGlnGlnGluProGlyValGlnGlyProLeuLysValGlyAsnSerLeu  231

```



```

Db      222 ProtrpalaSerProGlnThrMetGlnArgLeuSerArgLeuTyrAspPheSerPheArg 241
Qy      823 CAATTTTATGCTACAAACGACGCGGAGTTCGCCGCGAGCGCGCCACCCGCTATTG 882
Db      242 PheLeuPheGlyIleTyrGlnGlnAlaGluLysAlaArgLeuGlnGlyGlyValLeuLeu 261
Qy      883 GATTGTATCATGGCAGCGGTGAACGCCCATCCACCGCAAAACAGGCGGTATGTGTGACA 942
Db      262 AlaGlnIleArgLysAsnLeuThrLeuMetAlaThrThrSerGln----- 276
Qy      943 TTACCCACTTCAGTACTGTTTATTGCGGACACATATACTGCGCAATCTCGCGCGG 1002
Db      277 LeuPro-----LysLeuLeuValTyrSerAlaHisAspThrThrLeuValAlaLeuGlnMet 295
Qy      1003 GCACTGGAGTCAACTGAGCGCTTCCCGGTGAGCGGATACACCGCGCGAGGTGTGAA 1062
Db      296 AlaLeuAspValTyr-----AsnGlyGlnGlnAlaProTyrAlaSerCys 310
Qy      1063 CTGGTGTGTTGAACGCTGGCGTAAAGCGATTAACAGCAGTGATTCAGGTTTCGCTG 1122
Db      311 HisIlePheGluLeuTyrGlnGluAspSerGlyAsn-----PheSerValGluMet 327
Qy      1123 GTCTTCAGACTTACAGCAGATCGGTGATAAACGCGGCTGTCTATTAAATACGCCGCC 1182
Db      328 TyrPhe-----ArgAsnGluSerAspLysAlaProTyrProLeuSerLeuPro--- 343
Qy      1183 GGAGAGGTGAACGCTGCGGAGGATGTGAAGAGCGAAATGCGCGGCGATGTGTTCG 1242
Db      344 -----GlyCysProHisArg-----CysPro 350
Qy      1243 TTGCAGGTTTACGCAATCGTAATGAAGCAGCATACCG-----GCG 1287
Db      351 LeuGlnAspPheLeuArgLeu---ThrGluProValValProLysAspTyrGlnGlnGlu 369
Qy      1288 TGCAGTTTGAGATCT 1302
Db      370 CysGlnLeuAlaSer 374

RESULT 12
T16058
hypothetical protein F13D11.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C;Accession: T16058
R;Rulton, L.
submitted to the EMBL Data Library, November 1995
A;Description: The sequence of C. elegans cosmid F13D11.
A;Reference number: S69020
A;Accession: T16058
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-416 <FUL>
A;Cross-references: UNIPROT:Q19390; EMBL:U40939; NID:g1072175; PID:g1072178; PIDN:AAA817
A;Genetics:
C;Superfamily: mammalian acid phosphatase

Alignment Scores:
Pred. No.: 0.00968 Length: 416
Score: 133.00 Matches: 104
Percent Similarity: 36.94% Conservative: 60
Best Local Similarity: 23.42% Mismatches: 174
Query Match: 5.66% Indels: 106
DB: 2 Gaps: 24

US-10-034-985-1 (1-1323) x T16058 (1-416)

Qy      1 ATGAAGCGATCTTAATCCCAATTTTATCTCTCTGATTTCGTTAACCCGCAATCTGCA 60
Db      1 MetArgLeuValLeuLeuPhe-----PheLeuPheProVal-----Ala 14

```

```

Qy      61 TTCGCTCAGAGTCAGCGGAGCTGAAGCTGGAAGTGTGTGTGATTGTGATGTCAGTCGTCATGGT 120
Db      15 PheGly-----LysLeuPheValGlnIleTyrPheArgHisGly 28
Qy      121. GTGGGTCTCCAAACCAAGCCACGCAACTGATGAGGATGTCAACCCAGAC----- 171
Db      29 GluArgThrProGlyHisTyrLeuTyrPheProGlyAspAspLeuAsnValAspTyr 48
Qy      172 -----GCATGGCCCAACTGCGCGGTAAATCTGGTTGGCTGACACCGCGNGGTGGT 222
Db      49 GlnGlnIleAlaTyrPro-----GlyGluLeuThrLysArgGlyLe 62
Qy      223 GAGCTAATCCCTATCTCGGACATATACAAAGCGAGGCTCTGTGTAGCCGAGGATTCTGT 282
Db      63 LeuGluGluPheGlnLeuGly-----GlnArgLeuArgLysIleTyrGlyGluHisPhe 80
Qy      283 GCGAAAAAGGCTGCCCGCAGTCTGGTTCAGTTCGCGATTATTGCTGATGTCGAGCAGCGT 342
Db      81 GlyAspThrTyrGlnProArgAspPheHisVal-----TyrThrGlyLysAspAsnArg 98
Qy      343 ACCCGTAAACAGCGCAAGCTTTCGCGCGCGGCTGGCACCTGACTGTGCAATAACCGTA 402
Db      99 ThrSerAlaSerAlaGlnAlaMetPheAlaGlyPheLeuProPro----- 113
Qy      403 CATACCCAGCAGATAGTCCAGTCCGATCCGCTATTATTAACTCTCTAAAACATGGCGTT 462
Db      114 ---AsnGluAspGlnThrTrpAsnTyrGluLeuLysTrpGlnPro-----Val 128
Qy      463 TGCCTAATCGATAACGCAAGCTGACTGACGCGATCTCAGCAGGCGGAGGAGGTCAT 522
Db      129 AlaGlnLeuThrAspGluSerIleAspTyrValSerLeu-----GlyAlaIle 144
Qy      523 -----GCTGACTTTCGCGGCAATCGGCAACCGCGTTTCGCGAACTGGACCGGTCT 576
Db      145 AspAsnCysProValTyrGlyGluAlaGlnArgLysSerSerGluTyrAlaGluValMet 164
Qy      577 AATTTCCGCAATCAAACTTGTGCTTAAAGCTGAGAAACAGACGAAAGCTGTTCATTA 636
Db      165 Asp-----GlnMetGluLysTyrAsp-----AlaGluLeu 174
Qy      637 ACCGAGCATACCATCGGNACTCAAGTGAGCGCGCAATGTCATCATTAACCGGTGCG 696
Db      175 LeuGlnLeuValArgAsnHis-----AlaAspGlu---ProIleValGluAla 189
Qy      697 GTAAGCCTCGCATCAATGCTGCGGAGATA-----TTTCTCTGCAACAGCACAG 747
Db      190 ValLysTyrAsnHisValIleAspSerLeuLysValArgTyrIleLeuGlnAspArg 209
Qy      748 GGAATCCGAGCGCGGGTGGGAAAGATCACCGATTACACAGGTGGAACACCTTGCTA 807
Db      210 ---LeuProTyrProGluTyrAlaArgGlyTyrGluAsnArgIleLeuAsnMetSerPhe 228
Qy      808 AGTTTGCATAACCGCAATTTTATTGCTACACGCGCAGGAGTTCGCCGAGCGCGC 867
Db      229 LeuIleHisAspAla-----ValValLysValGln 238
Qy      868 GCCACCCCGTATTGGAT-----TTGATCATGCGCGGTTGACGCCCATCCA 915
Db      239 AsnAspSerValGlyAspTyrHisAsnGluLeuValMetSerTyrPheGluThrHisLeu 258
Qy      916 CCGCAAAAAACAGCGGTATGTTGTCACATTATCCCACTTCAGTACTGTTTATTCGCGGACAC 975
Db      259 GlnLysAsnSerThrLysGlyVal-----PheIleSerGlyHis 271
Qy      976 GATACATACTCGCAATCTCGCGGCGGCACTGAGTCAACTGAGGCTTCCCGGTGAG 1035
Db      272 AspThrAsnLeuValThrIleTyrGluSerLeuArgLeuAsp-----GlyHis 287
Qy      1036 CCGGATAACACCGGCCA---GGTGTGAACTGTGTTCACACGCTGCGGTGCGGTAAGC 1092
Db      288 ProGluAspIleProAsnTyrGlyAlaHisIleAlaIleGlu----- 301
Qy      1093 GATAACAGCCAGTGGATTTCAGGTTTCGCTGCTCTCCAGACTTTTACAGCAGATGCGTGAT 1152

```

```
Db 302 -----MetHisGluProValGlyGlnLeuSerIle 311
QY 1153 AARAGCGGCTGTCATTAAATACGCCCGCGAGAGGTGAACCTGACCCCTGGCAGGATGT 1212
Db 312 LysPhePheLeuSerMetGlyPheAsnGlnThrArgValGluLeuHisProHisPheCys 331
QY 1213 GAAGAGCAATCGCAGGCGCATGTCGTGGCAGGTTTACGCAATCGTGAATGAA 1272
Db 332 ---SerArgSerGlnAsnAsnAspCysThrTrpAspGluPheGlnArgLeuValLysLys 350
QY 1273 GCACGCATACCG 1284
Db 351 SerArgLysPro 354

RESULT 13
A33395
acid phosphatase (EC 3.1.3.2) precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 29-Jan-1990 #sequence_revision 29-Jan-1990 #text_change 09-Jul-2004
C:Accession: A33395
R:Himeno, M.; Fujita, H.; Noguchi, Y.; Kono, A.; Kato, K.
Biochem. Biophys. Res. Commun. 162, 1044-1053, 1989
A>Title: Isolation and sequencing of a cDNA clone encoding acid phosphatase in rat liver
A:Reference number: A33395; MUID:89350910; PMID:2764916
A:Accession: A33395
A>Status: preliminary
A:Molecule type: mRNA
A:Cross-references: 1-423 <HIM>
A:Residues: 1-423 <HIM>
C:Superfamily: mammalian acid phosphatase
C:Keywords: phosphohistidine; phosphoprotein; phosphoric monoester hydrolase
F:41/Active site: Arg #status predicted
F:42/Active site: His (phosphohistidine intermediate) #status predicted

Alignment Scores:
Pred. No.: 0.0141 Length: 423
Score: 131.00 Matches: 105
Percent Similarity: 33.48% Conservative: 49
Best Local Similarity: 22.83% Mismatches: 184
Query Match: 5.58% Indels: 122
DB: 2 Gaps: 23

US-10-034-985-1 (1-1323) x A33395 (1-423)
QY 4 AAAGCATCTTAATCCATTTT-----TCTCTTCTGATTCGTTAAACCCGCAATCT 57
Db 10 GlnAlaAlaLeuGlnPheLeuLeuGlyMetCysLeuMetValMetProIleGln 29
QY 58 GCATTCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGGTGTGTCAGTCGTCAT 117
Db 30 Ala-----ArgSerLeuArgPheValThrLeuLeuTyrArgHis 42
QY 118 GGTGTGCGTGTCCAAAGCCAGCCAACTGATGAGGATGTCACCCAGAGCATGG 177
Db 43 GlyAspArgSerProValLysAla-----TyrProLysAspProTyrGlnGlu 58
QY 178 CCAACCTGGCGGTAAACCTGGTGTGCTGACACCGCGGGTGTGAGCTAATCGCCAT 237
Db 59 GluLysTrpProGlnGlyPheGlyGlnLeuThrLysGluGlyMetLeuGlnHisTrpGlu 78
QY 238 CTGGACATTACCAACGCGAGCTGTGTGTCGCCAGCATGTCGTGGCGAAAGGGGTGC 297
Db 79 LeuGlyGlnAlaLeuArgGlnArgTyr-----HisGlyPheLeu-----AsnAla 93
QY 298 CCGCAGTCTGTCAGGTGCGGATTAATTCGTGANTCGAC---GAGCGTACCCGTAAACA 354
Db 94 SerTyrHisArgGlnValTyrValArgSerThrAspPheAspArgThrLeuMetSer 113
QY 355 GGCGAAGCTTCGCGCGCGGCTGGCACCCTGCTGTGCAATAACCGTACATACCCAGCA 414
Db 114 AlaGluAlaAsnLeuAlaGlyLeuPheProPro-----ThrGluVal 127
```

```
QY 415 GATACGTCAGTCCGATCCGATCGTTATTATTAATCTCTTAAATACTGGCGTTTGCACACTGGAT 474
Db 128 GlnHisPheAsnProAsnIleSerTrpGlnProIle-----ProValHis 142
QY 475 AACGCGAACTGACTACGCGATCTCAGCGATCTCAGCAGGCGAGGGTCAATTCGTGACTTTACC 534
Db 143 ThrValProIleThrGlu-----ThrArgGlnThrProGlu 174
QY 535 GGCGCATCGGCAACCGGTTTCGGAACCTGGAACGGTCTTAATTTTCGCAATCAAC 594
Db 149 -----AspArgLeuLeuLysPheProLeuGlyPro 158
QY 595 TTGTGCTTAAACGTGAGAACAGCAGCAAGCTGTTCATTAACGAGGCAATCACCATCG 654
Db 159 CysProArgTyrGluGlnLeuGlnAsnGlu-----ThrArgGlnThrProGlu 174
QY 655 GAACTCAAGTGAAGCGCGCAATGCTCTCATTAACCGGTGCGGTAAGCCTCGCATCATG 714
Db 175 TyrGlnAsnMetSerIleGlnAsnAlaGlnPheLeuAspMetValAlaAsnGluThrGly 194
QY 715 CTGACGGAGATATTCTC-----CTGCAACAA 741
Db 195 LeuMetAsnLeuThrLeuGluThrIleTrpAsnValTyrAspThrLeuPheCysGluGln 214
QY 742 GCACAGGATGCGGAGCGCGGTGGGAAGGATCACCGATTTCACACCATGGAACACC 801
Db 215 ThrHisGlyLeuLeuProProTrp-----AlaSerProGlnThrValGlnAla 231
QY 802 TTGCTAAGTTTGCATTAACGGCAATTT-----TATTGCTACAACGCGCAGGAGTTGCC 858
Db 232 LeuSerGlnLeuLysAspPheSerPheLeuPheGlyIleHisAspGlnValGln 251
QY 859 CGCAGCGC-----GCCACCCGTTATTGGATTGATTCATCGGCA 897
Db 252 LysAlaArgLeuGlnGlyValLeuLeuAlaGlnLeuLysAsnLeuThrLeuMet 271
QY 898 GCGTTGACGCGCCATCCACGCAAAACAGCGCGTATGTTGTGATTCACCATTCAGTA 957
Db 272 AlaThrThrSerGlnPheProLys-----Leu 280
QY 958 CTGTTTATTCGCGGACACGATCTACTATCTGCGAAATCTCGCGCGGCGACTGGAGCTCAAC 1017
Db 281 LeuValTyrSerAlaHisAspThrThrLeuValAlaLeuGlnMetAlaLeuAsnValTyr 300
QY 1018 TGGACGCTTCCTCGGTACGCGGATAACACGCCCGCAGGTGGTGAACCTGGTGTGAACGC 1077
Db 301 -----AsnGlyLysGlnAlaProTyrAlaSerCysHisIlePheGluLeu 315
QY 1078 TGGCGTGGCTAAGCGATAACAGCCAGTGGATTTCAGGTTTCGTGCTTCCAGACTTTA 1137
Db 316 TyrGln-----GluAspAsnGlyAsnPhe---SerValGluMetTyrPheArgAsnAsp 332
QY 1138 CACGAGTGGCTGATAAAACCGCGCTGTCTATTAAATACGCCGCGGAGGAGGTGAACCTG 1197
Db 333 SerLys-----LysAlaProTrp-----ProLeu 340
QY 1198 ACCTGTCAGATGTGAAGAGCGAAATGCGCGGCGATGTGTTGTTGCGAGGTTTACG 1257
Db 341 ThrLeuProGlyCysProHisArg-----CysProLeuGlnAspPheLeu 355
QY 1258 CAAATCGTGAATGAAGCAGCATACCG-----CGTGCATTTGAGATCT 1302
Db 356 ArgLeu---ThrGluProValIleProLysAspTrpGlnLysGluCysGlnLeuAlaSer 374

RESULT 14
S14742
acid phosphatase (EC 3.1.3.2) precursor - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S14742
R:Geier, C.; von Figura, K.; Pohlmann, R.
Biochem. Biophys. Res. Commun. 199, 301-304, 1991
A>Title: Molecular cloning of the mouse lysosomal acid phosphatase.
```


A;Reference number: S14742; MUID:91282986; PMID:2059337

A;Accession: S14742

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-421 <GB>

A;Cross-references: UNIPROT:P24638; EMBL:X57199; NID:G52870; PID:CAA0485.1; PID:G52871

C;Superfamily: mammalian acid phosphatase

C;Keywords: lysosome; phosphoric monoester hydrolase

Alignment Scores:

Pred. No.:	0.0702	Length:	421
Score:	122.50	Matches:	97
Percent Similarity:	32.88%	Conservative:	49
Best Local Similarity:	21.85%	Mismatches:	191
Query Match:	5.21%	Indels:	107
DB:	2	Gaps:	19

US-10-034-985-1 (1-1323) x S14742 (1-421)

Qy	4	AAAGCGATCTTAATCCCATTT-----TTATCTCTTGATTCGGTTAACC	48
Db	8		
Qy	49	CCGCAATCTGCATTCGCTCAGAGTGAGCCGAGCTGAAGCTGGAAGTGGTGTATGTC	108
Db	28	AlaArgSer-----LeuArgPheValThrLeuLeu	37
Qy	109	AGTCGTATGTTGGTGGTCTCCAAAGCCACGCAACTGATGCAGGATGTCCACCCCA	168
Db	38	TyrArgHisGlyAspArgSerProValysThr-----TyrProLysAspPro	53
Qy	169	GACCATGCGCAACTCGCCGGTAAACTGGTGGTGTGACACCGCNGTGGTGAGCTA	228
Db	54	TyrGlnGluGluLysTrpProGlnGlyPheGlyGlnLeuThrLysGluGlyMetLeuGln	73
Qy	229	ATCGCCTATCTCGACATACCAACGCGAGGTGTGTAGCCGAGATGCTGGCGAAA	288
Db	74	HisTrpGluLeuGlyGlnAlaLeuArgGlnArgTyr-----HisGlyPheLeu-----	89
Qy	289	AAGGGCTGCCCGAGTCTGGTGGTGGTATTTGCTGATGTGCAGC---GAGCGTACC	345
Db	90	---AsnThrSerTyrHisArgGlnGluValTyrValArgSerThrAspPheAspArgThr	108
Qy	346	CGTAAACAGCGAAGCTTCGCCCGCGGTGACACTGCTGCAATAACCGTACAT	405
Db	109	LeuMetSerAlaGluAlaAsnLeuAlaGlyLeuPhePro-----	122
Qy	406	ACCAGGAGATAGTCCAGTCCGATCCGATCTTATTTATCTCTTAAACTGGCGTTGC	465
Db	123	AsnGluValGlnHisPheSerProAsnIleSerTrpGlnProIle-----	137
Qy	466	CAACTGGATACCGCAAGCTGACTGACGCGATCCTCAGCAGGCGAGGGTCAATTGCT	525
Db	138	ProValHisThrValProIleThrGlu-----	146
Qy	526	GACTTTACCGGGCATCGCAACCGCGTTTCGCAACTGGAAACGGGTCTTAATTTTCG	585
Db	147	-----AspArgLeuLeuPhePhePro	153
Qy	586	CAATCAAACTTGCTTAAAGTGTAGMAACAGACGAAAGCTTCTTAAACGACGCA	645
Db	154	LeuglyProCysProArgTyrGluGlnLeuGlnAsnGlu-----ThrArgGln	169
Qy	646	TTACCATCGGAACCAAGTGTGACGCGCGACATGCTCTATTAACTGGGTGAAGCTC	705
Db	170	ThrProGluTyrGlnAsnArgSerIleGlnAsnAlaGlnPheLeuAsnMetValAlaAsn	189
Qy	706	GCATCAATGCTGACGGAGATATTTCTC-----	732
Db	190	GluThrGlyLeuThrAsnValThrLeuGluThrIleTrpAsnValTyrAspThrLeuPhe	209
Qy	733	CTGCAACACAGCAGGGAATCCGGAGCGGGGTGGGAGGATCACCGATTCACACAG	792

Db	210	CysGluGlnThrHisGlyLeuLeuLeuProTrpAlaSerProGlnThrValGlnArg	229
Qy	793	TGGAACACCTTGTAAAGTTTGTCATAACGCGCAATTTTATTTGCTACAAACGCCAGAG	852
Db	230	LeuSerGlnLeuLysAspPheSerPheLeuPheLeuGlyIleHisGluGlnValGln	249
Qy	853	GTTCGCCGACGCGCCGCCCGTATTGATTTGATTCATGCGAGGTGAGCGCCCAT	912
Db	250	LysAlaArgLeuGlnGlyValLeuLeuAlaGlnIleLeuLysAsnLeuThrLeuMet	269
Qy	913	CCACCGCAAAACAGCGCGTATGCTGCACATTATCCCACTTCAGTACTGCTTATTTGCGCGA	972
Db	270	AlaThrThrSerGlnPhe-----ProLysLeuLeuValTyrSerAla	283
Qy	973	CACGATATAATCTGGCAATCTCGCGCGCAGCTGAGCTCACTGAGCGTTCGCCGT	1032
Db	284	HisAspThrThrLeuValAlaLeuGlnMetAlaAsnValTyr-----	298
Qy	1033	CAGCCGATACAGCGCGCGTGGTGAATCGTGTGTTCAACGCTGCGCTCGGCTAAGC	1092
Db	299	AsnGlyLysGlnAlaProTyrAlaSerCysHisIlePheGluLeuTyrGln-----Glu	316
Qy	1093	GATAACAGCCAGTGGATTTCAGTTCGCTGCTCTTCAGACTTTACAGCAGATCGGTAT	1152
Db	317	AspAsnGlyAsnPhe---SerValGluMetTyrPheArgAsnAspSerLys-----	332
Qy	1153	AAAACCCCGTGTCTTAAATACGCCCGCGAGAGGTGAACTGACCTCGGCGAGGATGT	1212
Db	333	LysAlaProTrpProLeuIleLeuPro-----GlyCys	343
Qy	1213	GAAGCGGAATCGCGAGGCGATGTTCTGTCGAGGTTTACGCAAACTCGTGAATGAA	1272
Db	344	ProHisArg-----CysProLeuGlnAspPheLeuArgLeu---ThrGlu	357
Qy	1273	GCAGCATACCG	1284
Db	358	ProValIlePro	361

RESULT 15

JH0610

acid phosphatase (EC 3.1.3.2) ACPP precursor [validated] - human

N;Alternate names: acid phosphatase, prostatic; orthophosphoric monoester phosphohydrolase

C;Species: Homo sapiens (man)

C;Date: 17-Aug-1992 #sequence, revision 01-Dec-1995 #text change 09-Jul-2004

C;Accession: JH0610; J0693; A38608; S01331; A32419; S1147; S38863; S41251; S17042; S42

R;Sharief, F.S.; Li, S.S.L.

Biochem. Biophys. Res. Commun. 184, 1468-1476, 1992

A;Title: Structure of human prostatic acid phosphatase gene.

A;Reference number: JH0610; MUID:92272747; PMID:1375464

A;Accession: JH0610

A;Molecule type: DNA

A;Residues: 1-386 <SHA>

A;Cross-references: UNIPROT:P15309; GB:M97580; GB:M97581; GB:M97582; GB:M97583; GB:M97588

A;Accession: J0693

A;Molecule type: mRNA

A;Residues: 1-386 <SH3>

A;Cross-references: GB:M97589; NID:G189611; PIDN:AAA60021.1; PID:G189613

R;Van Baten, R.L.; Davidson, R.; Stevis, P.E.; MacArthur, H.; Moore, D.L.

J. Biol. Chem. 266, 2313-2319, 1991

A;Title: Covalent structure, disulfide bonding, and identification of reactive surface

A;Reference number: A38608; MUID:9115848; PMID:1989985

A;Accession: A38608

A;Molecule type: mRNA

A;Residues: 1-386 <VAN>

A;Cross-references: GB:M34840; NID:G189620; PIDN:AAA69694.1; PID:G189621

A;Note: Part of this sequence, including the amino end of the mature protein, was confir

R;Vinko, P.; Virkkunen, P.; Henttu, P.; Koiko, K.; Solin, T.; Hunkala, M.L.

FEBS Lett. 236, 275-281, 1988

A;Title: Molecular cloning and sequence analysis of cDNA encoding human prostatic acid p

A;Reference number: S01331; MUID:88312981; PMID:2842184

A;Accession: S01331

A;Molecule type: mRNA

A;Residues: 1-14, 'A', '16', 'ASC', '20', 'CF', '23', 'C', '25-65', 'WTWTPHPA', '74-211', 'A', '213-386 <VIH>

A;Cross-references: EMBL:X52174; NID:g28321; PIDN:CAA36422.1; PID:g28322
A;Note: part of this sequence, including the amino end of the mature protein, was confirmed by Sharief, F.S.; Lee, H.; Leuderman, M.M.; Lundwall, A.; Deaven, L.L.; Lee, C.; Li, S.S. Biochem. Biophys. Res. Commun. 160, 79-86, 1989
A;Title: Human prostatic acid phosphatase: cDNA cloning, gene mapping and protein sequencing
A;Reference number: A32419; MUID:89228054; PMID:2712834
A;Accession: A32419
A;Molecule type: mRNA
A;Residues: 1-14, 'A', 16, 'ASC', 20, 'CF', 23, 'C', 25-94, 'D', 96-115, 'R', 117-214, 'S', 216-293, 'T'
A;Cross-references: GB:M42902; NID:g189618; PIDN:AA60022.1; PID:g189619
A;Note: the authors translated the codons GAC for residue 95 as Glu, CGT for residue 116 as Arg, and GGT for residue 117 as Gly.
R;Taylor, P.G.; Govindan, M.V.; Patel, P.C. Nucleic Acids Res. 18, 4928, 1990
A;Title: Nucleotide sequence of human prostatic acid phosphatase determined from a full-length cDNA
A;Reference number: S11147; MUID:90370491; PMID:2395659
A;Accession: S11147
A;Status: preliminary; translation not shown
A;Molecule type: mRNA
A;Residues: 1-14, 'AFASC', 20, 'CF', 23, 'C', 25-45, 'H', 47-65, 'RIMPTHPA', 74-138, 'E', 140-156, 'R'
A;Cross-references: EMBL:X53605; NID:g35683; PIDN:CAA37673.1; PID:g35684
R;Banas, B.; Blaschke, D.; Pittler, F.; Hoerz, W. submitted to the EMBL Data Library, April 1993
A;Description: Characterization of the promoter of the human prostatic acid phosphatase gene
A;Reference number: S38863
A;Accession: S38863
A;Molecule type: DNA
A;Residues: 1-40 <BAN>
A;Cross-references: EMBL:X71391; NID:g416530; PIDN:CAA50514.1; PID:g416531
R;Virkkunen, P.H.; Hedberg, P.; Palvimo, J.J.; Birt, E.; Porvari, K.; Taavitsainen, P.; submitted to the EMBL Data Library, September 1993
A;Description: Structural organization of human and rat prostate-specific acid phosphatase gene in the human gene promoter.
A;Reference number: S41251
A;Accession: S41251
A;Molecule type: DNA
A;Residues: 1-40 <VIR>
A;Cross-references: EMBL:X74961; NID:g439657; PIDN:CAA52913.1; PID:g439658
R;Lee, H.; Chu, T.M.; Li, S.S.L.; Lee, C. Biochem. J. 277, 759-765, 1991
A;Title: Homodimer and heterodimer subunits of human prostate acid phosphatase.
A;Reference number: S17042; MUID:91336999; PMID:1908222
A;Accession: S17042
A;Status: preliminary
A;Molecule type: protein
A;Residues: 33-49 <LEE>
R;Banas, B.; Blaschke, D.; Pittler, F.; Hoerz, W. Biochim. Biophys. Acta 1217, 188-194, 1994
A;Title: Analysis of the promoter of the human prostatic acid phosphatase gene.
A;Reference number: S42730; MUID:94153995; PMID:8110833
A;Accession: S42730
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-40 <BA2>
A;Cross-references: GB:X71391; NID:g416530; PIDN:CAA50514.1; PID:g416531
R;Sharief, F.S.; Li, S.S. Biochem. Mol. Biol. Int. 33, 561-565, 1994
A;Title: Nucleotide sequence of human prostatic acid phosphatase ACP gene.
A;Reference number: I37175; MUID:95038536; PMID:7951074
A;Accession: I37175
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-386 <RES>
A;Cross-references: EMBL:U07097; NID:g515995; PIDN:AA860640.1; PID:g515997
C;Comment: This protein is synthesized under androgen regulation by epithelial cells of C636 rat prostate carcinoma cells.
C;Genetics:
A;Gene: GDB:ACPP
A;Cross-references: GDB:119644; OMIM:171790
A;Map position: 3q21.3-3q25.2
A;Introns: 40/3; 72/3; 101/3; 152/3; 185/3; 216/3; 260/3; 288/3; 323/2
C;Function:
A;Description: catalyzes the hydrolysis of a wide range of phosphate esters
C;Superfamily: mammalian acid phosphatase
C;Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrolase

F;1-32/Domain: signal sequence #status predicted <SIG>
F;33-386/Product: acid phosphatase ACP #status experimental <MAT>
F;43, 86/Active site: Arg #status predicted
F;44/Active site: His (phosphohistidine intermediate) #status predicted
F;94, 220, 333/Binding site: carbonylate (Asn) (covalent) #status experimental
F;161-372, 215-313, 347-351/Disulfide bonds: #status experimental

Alignment Scores:
Pred. No.: 0.196 Length: 386
Score: 117.00 Matches: 87
Percent Similarity: 34.09% Conservative: 63
Best Local Similarity: 19.77% Mismatches: 178
Query Match: 4.98% Indels: 112
DB: 1 Gaps: 15

US-10-034-985-1 (1-1323) x JH0610 (1-386)

QY 7 GCGATCTTAATCCATTTTATCTTCTGATTCCTTAACCCGCAATTCGATTCGCT 66
DB 13 SerLeuSerLeuGlyPheLeuPheLeuPhePheTrpLeuAspArgSerValLeuAla 32
QY 67 CAGAGTGACCGGAGCTGAAGCTGGAAGTGTGTGATTCAGTGTCTGATGTCGCT 126
DB 33 Lys -----GluLeuLysPheValThrLeuValPheArgHisGlyAspArg 47
QY 127 GCTCCAAACCAAGCCAGCAACTGATGATGATGTCAACCCAGCAGCATGCCAACCC 183
DB 48 SerProle -----AspThrPheProThrAsp 56
QY 184 -----TGCCGGTAAACTGGTGTGGCTGACACCGCGNGTGTGTGAG 225
DB 57 ProileLysGluSerSerTrpProGlnGlyPheGlnLeuThrGlnLeuGlyMetGlu 76
QY 226 CTAATCGCTATCTCGGACATTAACCAAGCCAGCGTCTGTAGCCGAGGATTCGTCG 285
DB 77 GlnHisTrpGluLeuGlyGluTrpIleArgLysArgTrp ----- 89
QY 286 AAAAAGGGTCCCGCAGTCT -----GGTCAGTTCGCGATTATTGCTGATGTCGAC 336
DB 90 ArgLysPheLeuAsnGluSerTrpLysHisGluInValTrpIleArgSerThrAspVal 109
QY 337 GAGCTACCGGTAAACACAGCGGAAGCTTCGCCCGGGGTGGCAGCTGACTGTGCNATA 396
DB 110 AspArgThrLeuMetSerAlaMetThrAsnLeuAlaLeuPheProGluGlyVal 129
QY 397 ACCTATACATCCAGGACAGATCGTCCAGTCCCGTCCGTTATTATCTCTAAAACCT 456
DB 130 SerIleTrp -----AsnProileLeuLeuTrpGlnProleProVal 143
QY 457 GCGGTTTCCCACTGGATACGGAACGTGACTGACCGCATCTCAGCAGGCGAGGAGG 516
DB 144 HisThrValProleuSerGlu ----- 150
QY 517 TCAATTGCTGACTTTTACCGGGCATTCGGAACCGCGTTTCGGAACCGGTCTT 576
DB 151 -----AspGlnLeuLeu 154
QY 577 AATTTCCGCAATCAACTTGTGCTTAAACGTGAGAAACAGGACGAAAGCTGTTCTTA 636
DB 155 TyrLeuProPheArgAsn ---CysProArgPheGlnGluLeuGluSerGluThrLeuLys 173
QY 637 ACCGAGCATTTACCATCGGAACCTCAAGGTGAGCGCGCACAAATCTCATTAACCGGTGCG 696
DB 174 SerGluGluPheGlnLysArgLeuHisProTyrLysAspPheIleAlaThrLeuGlyLys 193
QY 697 GTAAGC -----CTCGCATCAATGCTGAGGAGATATTT 729
DB 194 LeuSerGlyLeuHisGlyGlnAspLeuPheGlyIleTrpSerLysValTyrAspProLeu 213
QY 730 CTCTTCACACAGCAGGGAATGCGGAGCGGGTGGGAGGAGGATTCACCGATTCACAC 789
DB 214 TyrCysGluSerValHisAsnPheThrLeuProSerTrpAlaThrGluAspThrMetThr 233

THIS PAGE BLANK

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: March 11, 2005, 23:09:30 ; Search time 254 Seconds
(without alignments)
5334.494 Million cell updates/sec

Title: US-10-034-985-1
Perfect score: 2349
Sequence: 1 atgaagcgatcttaatccc.....atcacctaccatcactaa 1323

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Zgapop 6.0 , Zgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 3224756

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp
-Q=/cgn2 1/USPRO spool_p/US10034985/runat 10032005 164240 11950/app query.fasta_1.1479
-DB=uniprot_03 -OPMT=fastan -SUFFIX=rup -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MOD=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USR=US10034985 @CGN 1 1 271 @runat 10032005 164240 11950 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2235	95.1	432	1 PPA_ECOLI	P07102 escherichia
2	2231	95.0	432	2 O8GN88	O8gn88 escherichia
3	2229	94.9	432	2 O8RND6	O8rkd6 escherichia
4	2229	94.9	432	2 O8RND7	O8rkd7 escherichia
5	2229	94.9	432	2 O8RND8	O8rkd8 escherichia
6	2229	94.9	432	2 O8RKE0	O8rke0 escherichia
7	2228	94.8	432	2 O8RND9	O8rkd9 escherichia
8	2225	94.7	432	2 O8RND5	O8rkd5 escherichia
9	2222	94.6	442	2 O83RW2	O83rw2 shigella fl
10	2221	94.6	432	2 O6RK08	O6rk08 escherichia
11	2218	94.4	432	2 O7UD08	O7ud08 shigella fl
12	2213	94.2	446	2 O8CWF5	O8cw75 escherichia
13	2210	94.1	434	2 O7AFW3	O7afw3 escherichia
14	2210	94.1	444	2 O8XC29	O8xc29 escherichia
15	1374.5	58.5	433	2 O676V7	O676v7 citrobacter
16	1077	45.8	444	2 Q6U677	Q6u677 obesumbacte

17	1054.5	44.9	444	2	O6TQA8	O6taq8 obesumbacte
18	943.5	40.2	441	2	O669R3	O669r3 yersinia ps
19	943.5	40.2	441	2	O8ZFP6	O8zfp6 yersinia pe
20	545	23.2	413	2	O8XBZ6	O8xbz6 escherichia
21	541	23.0	417	1	AGP_PRORE	P19926 escherichia
22	539	22.9	413	1	AGP_ECOLI	O52309 providencia
23	538	22.9	413	2	O7UD02	O7ud02 shigella fl
24	538	22.9	421	2	O83RV6	O83rv6 shigella fl
25	529	22.5	413	1	AGP_SALTY	O33921 salmonella
26	526	22.4	414	2	O9AAQ4	O9aaq4 caulobacter
27	524	22.3	413	2	O8Z7P1	O8z7p1 salmonella
28	523	22.3	413	2	O6EV19	O6ev19 enterobacte
29	505	21.5	392	2	O8PP76	O8pp76 xanthomonas
30	505	21.5	443	2	O8P330	O8p330 xanthomonas
31	480.5	20.5	435	2	O8PF53	O8pf53 xanthomonas
32	467.5	19.9	435	2	O6CZF4	O6czf4 erwinia car
33	433.5	18.5	428	2	O8GD20	O8gd20 pseudomonas
34	392	16.7	318	2	O8CW72	O8cw72 escherichia
35	373	15.9	414	2	O8PPX6	O8ppx6 xanthomonas
36	359	15.3	421	2	O84CN9	O84cn9 klebsiella
37	353	15.0	421	2	O7WSY1	O7wsy1 klebsiella
38	233	9.9	419	2	O8VQS2	O8vq2 klebsiella
39	195	8.3	426	2	O9GNZ3	O9gnz3 leishmania
40	150	6.4	531	2	O9GNZ3	O9gnz3 leishmania
41	140	6.0	423	1	PPAL_HUMAN	P11117 homo sapien
42	140	6.0	542	2	O9UIA2	O9ui2 leishmania
43	139.5	5.9	333	2	O9BZG3	O9bzg3 homo sapien
44	139	5.9	462	2	O7QC05	O7qc05 anopheles g
45	137.5	5.9	406	2	Q66L09	Q66l09 xenopus lae

ALIGNMENTS

RESULT 1

PPA_ECOLI
ID PPA_ECOLI STANDARD; PRT; 432 AA.
AC P07102;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Periplasmic appA protein precursor [Includes: Phosphoanhydride
DE phosphorylase (EC 3.1.3.2) (pH 2.5 acid phosphatase) (AP); 4-
DE phytase (EC 3.1.3.26)].
GN Name:appA; OrderedLocusNames=b0980;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RX MEDLINE=90368616; PubMed=2168385;
RC STRAIN=K12;
RA Dassa J., Marck C., Boquet P.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase."
RL J. Bacteriol. 172:5497-5500(1990).
RN [2]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-29.
RC STRAIN=K12;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glaesner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shaoy Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glaesner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shaoy Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,

RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
RA Yano M., Horiuchi T.,
RT "A 718-kb DNA sequence of the *Escherichia coli* K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
RN [4]
RP SEQUENCE OF 1-112 FROM N.A.
RX MEDLINE=8721766; PubMed=3038201; DOI=10.1016/0300-9084(87)90045-9;
RA Touati E., Danchin A.;
RT "The structure of the promoter and amino terminal region of the pH 2.5
RT acid phosphatase structural gene (appA) of *E. coli*: a negative control
RT of transcription mediated by cyclic AMP.";
RL Biochimie 69:215-221(1987).
RN [5]
RP SEQUENCE OF 1-17 FROM N.A.
RX STRAIN=KL12;
RA MEDLINE=92049231; PubMed=1658595;
RA Dassa J., Feihi H., Marck C., Dion M., Kieffer-Bontemps M.,
RA Boquet P.L.;
RT "A new oxygen-regulated operon in *Escherichia coli* comprises the genes
RT for a putative third cytochrome oxidase and for pH 2.5 acid
RT phosphatase (appA).";
RL Mol. Gen. Genet. 229:341-352(1991).
RN [6]
RP CHARACTERIZATION, AND SEQUENCE OF 23-34.
RA Greiner R., Jany K.D.;
RT "Characterization of a phytase from *Escherichia coli*.";
RL Biol. Chem. Hoppe-Seyler 372:664-665(1991).
RN [7]
RP CHARACTERIZATION, AND SEQUENCE OF 23-35.
RX MEDLINE=93256556; PubMed=8387749; DOI=10.1006/abbi.1993.1261;
RA Greiner R., Konietzny U., Jany K.D.;
RT "Purification and characterization of two phytases from *Escherichia coli*.";
RL Arch. Biochem. Biophys. 303:107-113(1993).
RN [8]
RP CHARACTERIZATION.
RX MEDLINE=20161462; PubMed=10696472;
RA Glovan S., Wang G., Zhang J., Forsberg C.W.;
RT "Characterization and overproduction of the *Escherichia coli* appA
RT encoded bifunctional enzyme that exhibits both phytase and acid
RT phosphatase activities.";
RL Can. J. Microbiol. 46:59-71(2000).
RN [9]
RP MUTAGENESIS.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostalin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.-M.,
RA van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT *Escherichia coli* acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836(1992).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).
RX MEDLINE=20122624; PubMed=10655611; DOI=10.1038/72371;
RA Lam D., Golovan S., Forsberg C.W., Jia Z.;
RT "Crystal structures of *Escherichia coli* phytase and its complex with
RT pyruvate.";
RL Nat. Struct. Biol. 7:108-113(2000).
CC -1- CATALYTIC ACTIVITY: A phosphate monoester + H(2)O = an alcohol +
CC phosphate.
CC -1- CATALYTIC ACTIVITY: Myo-inositol hexakisphosphate + H(2)O = 1D-
CC myo-inositol 1,2,3,4,5-pentakisphosphate + phosphate.
CC -1- SUBUNIT: Monomer.
CC -1- SUBCELLULAR LOCATION: Periplasmic.
CC -1- INDUCTION: In addition to CAMP-mediated control, this enzyme is
CC induced when bacterial cultures reach stationary phase; its
CC synthesis is triggered by phosphate starvation or a shift from
CC aerobic to anaerobic conditions.
CC -1- SIMILARITY: Belongs to the histidine acid phosphatase family.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
CC or send an email to license@1sb-sib.ch).
CC -----
CC DR EMBL; M58708; AAA72086.1; -;
CC DR EMBL; U00096; AAC74065.1; -;
CC DR EMBL; D90735; BAA35745.1; -;
CC DR EMBL; X05471; BAA29031.1; -;
CC DR EMBL; M58708; -; NOT_ANNOTATED_CDS.
CC DR EMBL; S63811; AAB20286.1; -;
CC DR PIR; B36733; B36733.
CC DR PDB; 1DKL; X-ray; A/B=23-432.
CC DR PDB; 1DKM; X-ray; A=23-432.
CC DR PDB; 1DKN; X-ray; A=23-432.
CC DR PDB; 1DKO; X-ray; A=23-432.
CC DR PDB; 1DKP; X-ray; A=23-432.
CC DR PDB; 1DKQ; X-ray; A=23-432.
CC DR SWISS-2DPAGE; P07102; COLL.
CC DR ECODBASE; E80047; -;
CC DR EcoGene; EGI0049; appA.
CC DR InterPro; IPR000560; HisAc_phosphatse.
CC DR Pfam; PF00328; Acid_phosphat_A; 1.
CC DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
CC DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; 1.
CC KW 3D-structure; Complete proteome; Direct protein sequencing; Hydrolase;
CC MW Multifunctional enzyme; Periplasmic; Signal.
CC FT SIGNAL 1 22
CC FT CHAIN 23 432 Periplasmic appA protein.
CC FT ACT_SITE 39 39 Nucleophile.
CC FT ACT_SITE 326 326 Proton donor.
CC FT DISULFID 99 130
CC FT DISULFID 155 430
CC FT DISULFID 200 210
CC FT DISULFID 404 413
CC FT CONFLICT 51 66 MODVTPDAWPTWPKL -> NAGCHPRMANLAGKT (in
CC Ref. 3).
CC FT CONFLICT 75 76 EL -> DV (in Ref. 4).
CC FT CONFLICT 112 112 D -> S (in Ref. 4).
CC FT STRAND 28 38
CC FT STRAND 42 42
CC FT HELIX 49 53
CC FT TURN 54 54
CC FT TURN 55 67
CC FT STRAND 69 69
CC FT HELIX 71 90
CC FT TURN 91 92
CC FT TURN 102 103
CC FT STRAND 105 109
CC FT HELIX 113 126
CC FT TURN 128 129
CC FT STRAND 134 135
CC FT TURN 140 141
CC FT HELIX 145 147
CC FT TURN 149 153
CC FT HELIX 159 169
CC FT TURN 170 171
CC FT HELIX 174 179
CC FT TURN 180 181
CC FT HELIX 182 192
CC FT TURN 193 193
CC FT HELIX 194 196
CC FT TURN 198 201
CC FT TURN 202 202
CC FT HELIX 212 215
CC FT STRAND 220 223
CC FT TURN 224 225
CC FT STRAND 226 229
CC FT HELIX 231 249
CC FT TURN 250 250
CC FT HELIX 254 257

FT TURN 258 258
 FT HELIX 262 279
 FT TURN 280 280
 FT HELIX 283 301
 FT STRAND 307 308
 FT HELIX 310 312
 FT STRAND 314 315
 FT STRAND 319 324
 FT HELIX 326 336
 FT TURN 337 337
 FT STRAND 342 342
 FT TURN 343 344
 FT STRAND 345 345
 FT TURN 351 352
 FT STRAND 354 362
 FT TURN 363 366

Alignment Scores:
 Pred. No.: 2,31e-172 Length: 432
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 95.15% Indels: 0
 DB: 1 Gaps: 0

US-10-034-985-1 (1-1323) x PPA_ECOLI (1-432)

QY 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGGTTAAACCCCGCAATCTGCA 60
 DB 1 MetLysAlaIleuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
 QY 61 TTCGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGTGTGATGTGTCATGCTCATGTT 120
 DB 21 PheAlaGlnSerGluProGluLeuLeuLeuSerValValIleValSerArgHisGly 40
 QY 121 GTGCGTGTCCACACGAGCCAGCACTGATCAGGATGTCAACCCAGACGATGGCCA 180
 DB 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
 QY 181 ACCTGCGCGGTAAACTGGGTTGGCTGACACCGCGNGGTGGTGAGCTAATCGCTATCTC 240
 DB 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
 QY 241 GGCATATTACCAACGCGAGCTCTGTGTAGCCGACGAGTGTGCGCGAAAAGGGCTGCCG 300
 DB 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 QY 301 CAGTCTGTTCGCGGATTTATGCTGATGTCGACGAGCGGTACCGTAAACAGCGGAA 360
 DB 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 QY 361 GCCTTCGCGCGCGGTGGCACCTGACTGTGCAATAACCGTACATACCCAGCGAGATACG 420
 DB 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 QY 421 TCCAGTCCCGATCCGTTATTAATCTCTAAAAAAGCTGGGTTTGGCAACTGGATAACGCG 480
 DB 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 QY 481 AACGTGACTCAGCGATCTCAGCAGGGGAGGGGTCAATTCTGACTTTACCGGCGAT 540
 DB 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 QY 541 CGGCAAAAGCGGTTTCGCGAACTGGAAACGGGTCTTAATTTTCCGCAATCAAACTTGTGC 600
 DB 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 QY 601 CTTAAACGTGAGAAACAGGACGAAGCTGTTCATTAAACGAGGCAATTACCATCGGAATC 660
 DB 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 QY 661 AGGTGAGCGCGGACAAATGTCTCATTAACCGGTGGGTAAAGCCTCGCATCAATGTGACG 720

Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 QY 721 GAGATATTTCTCTGCAACAAGCACAGGAATGCGGAGCGGGTGGGAAGATCACC 780
 Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 QY 781 GATTACACAGTGAACACCTTGTAAAGTTTGCATAACCGCAATTTATTTGCTACAA 840
 Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
 QY 841 CGCAGCCAGAGGTTCGCCGCGCCGCGCCGCTTATTTGATTTGATTCATGCGGACG 900
 Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
 QY 901 TTGACGCCCCATCCACCCGCAAAACAGGCGTATGTTGTGACATTACCCACTTCAGTACTG 960
 Db 301 LeuThrProHisProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
 QY 961 TTTATTGCGGACACGATACTAATCTGGCAAAATCTCGGCGGCGCACTGGAGCTCAACTGG 1020
 Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 QY 1021 ACGTTTCCCGCTCAGCGGATACACCGCCGAGGTGGTGAACCTGTTGTTGAACGCTGG 1080
 Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 QY 1081 CGTCGCTGAAGCGATAACAGCCAGTGGATTTCAGTTTCGCTGCTTCCAGACTTTACAG 1140
 Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 QY 1141 CAGATCGGTGATAAAGCGCGCTGTCTATTAAATACGCCGCCGAGAGGTGAAACTGACC 1200
 Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 QY 1201 CTGGCAGGATGTGAAGACGCAATGCGAGGCGATGTTGTTGTCGAGGTTTACGCAA 1260
 Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 ATCGTGAATGAACACGACATACCGCGCTGCAGTTTG 1296
 Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

RESULT 2
 Q8GN88 PRELIMINARY; PRT; 432 AA.
 ID Q8GN88;
 AC Q8GN88;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE AppA.
 GN Name=appA;
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang Z., He J., Yao B., Zhou Y., Chen Y., Yi Y.;
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Chen Y., Zhu Z., Zhang Z., He J.;
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF537219; AAN28334.1; --
 DR HSSP; P07102; 1DKM.
 DR GO; GO:0003993; Fractid phosphatase activity; IEA.
 DR InterPro; IPR000560; Hisac_phsphtse.
 DR Pfam; PF00328; Acid_phosphat_A; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN 1.
 SQ SEQUENCE 432 AA; 47055 MW; 5B355D76E7377737 CRC64;

Alignment Scores:

Pred. No.:	4,88e-172	Length:	432
Score:	2231.00	Matches:	429
Percent Similarity:	99.54%	Conservative:	1
Best Local Similarity:	99.31%	Mismatches:	2
Query Match:	94.98%	Indels:	0
DB:	2	Gaps:	0

US-10-034-985-1 (1-1323) x Q8GN88 (1-432)

QY	1	ATGAAGAGCATCTTAATCCATTTTATCTCTTCTGATTCGGTTAAACCCGCAATCTGCA	60
DB	1	MetLysAlaIleuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla	20
QY	61	TTTCGTACAGTACGCGGAGCTGAAGCTGGAAGTGGTGAATGTCATGTCAGTCATGGT	120
DB	21	PheAlaGlnSerGluProGluLeuIleuLeuGluSerValIleValSerArgHisGly	40
QY	121	GTGCGTCTCCAAACGACGCAACTGATGACGAGGATGTCAACCCAGACGCAATGGCCA	180
DB	41	ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro	60
QY	181	ACCTGCGCGGTAAACTGGTGGTGTGACACCGCGNGTGTGTGACTAATCGCTATCTC	240
DB	61	ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu	80
QY	241	GGACATTACCAAGCGCGTCTGTGACGCGGATTCGTGGCGAAAGGGCTGCCCG	300
DB	81	GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro	100
QY	301	CAGTCGTGTCAGCTCCGATATTGTCATGTGCGAGCGGTACCGGTAAACAGGCGAA	360
DB	101	GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu	120
QY	361	GCCTTCGCGCGGCGTGCACCTGACTGTGCAATAACGGTACATACCAGCGAGATACG	420
DB	121	AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
QY	421	TCAGTCCGATCCGTTATTATCTCTTAAACCTGGCGTTTGCACCAATCAACTTGGC	480
DB	141	SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla	160
QY	481	AACTGTACTGACCGCATCTCAGCAGGCGAGAGGTCGAATGTGACTTTACCGGGCAT	540
DB	161	AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis	180
QY	541	CGCAACGCGTTCGCAACTGGAACGGTGCTTAATTTCCGCAATCAACTTGGC	600
DB	181	ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnProGlnSerAsnLeuCys	200
QY	601	CTTAAACGTGAGAAACAGGACGAAAGCTCTTCATTAACGAGCATTAACCACTGCACTC	660
DB	201	LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu	220
QY	661	AAGGTGAGCCGCAATGTCATTAATTAACCGGTGGGTAAAGCTCGCATCAATGCTGACG	720
DB	221	LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr	240
QY	721	GAGATATTTCTCTGCAACAGCACAGGAATCCGCGAGCCGGGTGGGGAAGATCAACC	780
DB	241	LysIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr	260
QY	781	GATTACACAGTGGAAACCTTGTCTAATGTTGCTATTAACGCGCAATTTTATTTGCTCAA	840
DB	261	AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln	280
QY	841	CGACGCGCAGGTTCGCGCGACCGCGCACCCGTTATTCGATTCATGTCAGCG	900
DB	281	ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuLysThrAla	300
QY	901	TTGACGCGCCATCCACCGCAAAACAGGCGTATGGTGTGACATTACCCACTTCAGTACTG	960
DB	301	LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu	320

US-10-034-985-1 (1-1323) x Q8KDK6 (1-432)

Pred. No.:	7,09e-172	Length:	432
Score:	2229.00	Matches:	429
Percent Similarity:	99.31%	Conservative:	0
Best Local Similarity:	99.31%	Mismatches:	3
Query Match:	94.89%	Indels:	0
DB:	2	Gaps:	0

US-10-034-985-1 (1-1323) x Q8KDK6 (1-432)

QY	961	TTTATTTCGGACACGATACCTAATCTCGCAAAATCTCGCGCGCACCTGGAGCTCAACTGG	1020
DB <th>321</th> <th>PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp</th> <th>340</th>	321	PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp	340
QY <th>1021</th> <th>ACGCTTCCCGTTCAGCCCGGATAACACCGCCGACGTCGTCGTCGTCGTCGTCGTCGTCG</th> <th>1080</th>	1021	ACGCTTCCCGTTCAGCCCGGATAACACCGCCGACGTCGTCGTCGTCGTCGTCGTCGTCG	1080
DB <th>341</th> <th>ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp</th> <th>360</th>	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp	360
QY <th>1081</th> <th>CCTCGGCTAAGGATAACAGCCAGTCGATTCAGGTTTCGTCGTCGTCGTCGTCGTCGTCG</th> <th>1140</th>	1081	CCTCGGCTAAGGATAACAGCCAGTCGATTCAGGTTTCGTCGTCGTCGTCGTCGTCGTCG	1140
DB <th>361</th> <th>ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln</th> <th>380</th>	361	ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln	380
QY <th>1141</th> <th>CAGATCGCTGTATAAACCCGCTGTCTATTAATAATACGCGCCGAGAGGTGAACCTGACC</th> <th>1200</th>	1141	CAGATCGCTGTATAAACCCGCTGTCTATTAATAATACGCGCCGAGAGGTGAACCTGACC	1200
DB <th>381</th> <th>GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr</th> <th>400</th>	381	GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr	400
QY <th>1201</th> <th>CTGGCAGGATGTGAAGACGAAATGCGCAGGCGATGTTCGTCGTCGTCGTCGTCGTCGTCG</th> <th>1260</th>	1201	CTGGCAGGATGTGAAGACGAAATGCGCAGGCGATGTTCGTCGTCGTCGTCGTCGTCGTCG	1260
DB <th>401</th> <th>LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln</th> <th>420</th>	401	LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln	420
QY <th>1261</th> <th>ATCGTGAATGAGCAGCATACCGCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG</th> <th>1320</th>	1261	ATCGTGAATGAGCAGCATACCGCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG	1320
DB <th>421</th> <th>IleValAsnGluAlaArgIleProAlaCysSerLeu</th> <th>432</th>	421	IleValAsnGluAlaArgIleProAlaCysSerLeu	432

RESULT 3

ID	Q8KDK6	PRELIMINARY;	PRT;	432 AA.
AC	Q8KDK6;			
DT	01-JUN-2002 (TrEMBLrel. 21, Created)			
DT	01-JUN-2002 (TrEMBLrel. 21, Last sequence update)			
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)			
DE	Periplasmic phosphoanhydride phosphohydrolase.			
GN	Names: appA;			
OS	Escherichia coli.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Escherichia.			
OX	NCBI_TaxID=562;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90368616; PubMed=2168385;			

QY 1 ATGAACCGATCTTAATCCCAATTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60
 Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20

QY 61 TTCGCTCAGAGTACGCGGAGCTGAAGCTGGAAGTGTGTGATTGTCAGTCTCATGGT 120
 Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40

QY 121 GTGCGTCTCAACCAAGGCCACCAACTGATGACGAGTGTACCCACGACGATGGCCA 180
 Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60

QY 181 ACCTGGCCGTAAACATGGGTGTGCTCACCCGNGGTGGTGAAGTAATCGCTATCTC 240
 Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80

QY 241 GGACATTACCAACGCGAGCTGTGGTAGCCGACGAGTGTGGGAAAGGGGTGCGCG 300
 Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100

QY 301 CAGTCTCGTCAGGTCGCGATTATTGCTGATGTCACGAGGTACCCGTAAACACAGGCGAA 360
 Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluAlaThrArgLysThrGlyGlu 120

QY 361 GCCTTCGCCCGCGGTGGCACCTGACTGTGCAATAACCGTACATACCCAGGACAGTACG 420
 Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140

QY 421 TCCAGTCCCGATCGGTTATTTATCTCTTAAACCTGCGGTTGCGCAACTGGATAACGCG 480
 Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160

QY 481 AACGTGACTCAGCGATCTCCTCAGCAGGCGAGGAGTCAATTGCTGATTTTACCGGSCAT 540
 Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180

QY 541 CGGCACACGCGGTTTCGCAACTGGAAACGGGTGCTTAATTTTCGCAATCAAACTGTGTC 600
 Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200

QY 601 CTTAAACGTGAGAAACAGGAGAAAGCTGTTTCAATTAAACGAGGATTAACCATCGAACTC 660
 Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220

QY 661 AAGTGAGCCCGACATGCTCATTAACCGTGCCTGAGCTCGCATCAATCGCTGACG 720
 Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240

QY 721 GAGATATTTCTCTGCAACACAGCACAGGAAATGCGGAGCGCGGTGGGGAAGGATCACCC 780
 Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260

QY 781 GATTACACAGTGGAAACACTGTGTAAGTTTGATTAACCGCAATTTTATTGCTACAA 840
 Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280

QY 841 CGCAGCCGACAGTTGCGCGCAGCGCCACCGCTTATTGATTTGATGTCATGGCAGCG 900
 Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300

QY 901 TTGACGCCCATCCACCGCAAAACAGCGCTATGGTGTGATGACATTAACCATCTCAGTACTG 960
 Db 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320

QY 961 TTTATTGCCGACACGATATACTTGGCAAACTCGCGCGCGCAGCTGGAGCTCAACTGG 1020
 Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340

QY 1021 ACGTTCCCGGTACGCGGATACACGCCCGCAGGTGGTGAAGTGGTGTGTTGACCGTGG 1080
 Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360

QY 1081 CGTCGCTAAGCCGATACAGCCAGTCAGATTTCAGTTTCGCTTCAGACTTTACAG 1140
 Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380

QY 1141 CAGATCGGTGATAAACGCGCTGTCTATTAATACGCCGCCGAGAGGTGAAACTGACC 1200
 Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400

QY 1201 CTGCGCAGGATGTGAAGAGCGAAATGCGCAGGCGATGTTCTGTTGGCAGGTTTACGCAA 1260
 Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420

QY 1261 ATCGTGAATGAACACGACATACCGCGCTGCAGTTTG 1296
 Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

RESULT 4
 Q8RKD7 PRELIMINARY; PRT; 432 AA.

AC Q8RKD7
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Periplasmic phosphoanhydride phosphohydrolase.
 GN Name=appa;
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90368616; PubMed=2168385;
 RA Dassa J., Marck C., Boquet P.-L.L.;
 RT "The complete nucleotide sequence of the Escherichia coli gene appa
 RT reveals significant homology between pH 2.5 acid phosphatase and
 RT glucose-1-phosphatase.";
 RL J. Bacteriol. 172:5497-5500 (1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93054596; PubMed=1429631;
 RA Ostain K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,
 RA Van Etten R.L.;
 RT "Overexpression, site-directed mutagenesis, and mechanism of
 RT Escherichia coli acid phosphatase.";
 RL J. Biol. Chem. 267:22830-22836 (1992).
 DR EMBL; L03373; AAA00005.1; -;
 DR HSSP; P07102; 1DKM.
 DR GO; GO:0003993; F:acid phosphatase activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR InterPro; IPR000560; HisAc_phsphtse.
 DR Pfam; PF00328; Acid_phosphat_A; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase.
 SQ SEQUENCE 432 AA; 46971 MW; 755D5E4B1AD916A6 CRC64;

Alignment Scores:
 Pred. No.: 7, 09e-172 Length: 432
 Score: 2229.00 Matches: 429
 Percent Similarity: 99.31% Conservative: 0
 Best Local Similarity: 99.31% Mismatches: 3
 Query Match: 94.89% Indels: 0
 DB: 2 Gaps: 0

US-10-034-985-1 (1-1323) x Q8RKD7 (1-432)

QY 1 ATGAACCGATCTTAATCCCAATTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60
 Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20

QY 61 TTCGCTCAGAGTACGCGGAGCTGAAGCTGGAAGTGTGTGATTGTCAGTCTCATGGT 120
 Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40

QY 121 GTGCGTCTCAACCAAGGCCACCAACTGATGACGAGTGTACCCACGACGATGGCCA 180
 Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60

QY 181 ACCTGGCCGTAAACATGGGTGTGCTCACCCGNGGTGGTGAAGTAATCGCTATCTC 240
 Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80

QY 241 GGACATTACCAACGCGAGCTGTGGTAGCCGACGAGTGTGGGAAAGGGGTGCGCG 300
 Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100

QY 301 CAGTCTCGTCAGGTCGCGATTATTGCTGATGTCACGAGGTACCCGTAAACACAGGCGAA 360
 Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluAlaThrArgLysThrGlyGlu 120

QY 361 GCCTTCGCCCGCGGTGGCACCTGACTGTGCAATAACCGTACATACCCAGGACAGTACG 420
 Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140

QY 421 TCCAGTCCCGATCGGTTATTTATCTCTTAAACCTGCGGTTGCGCAACTGGATAACGCG 480
 Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160

QY 481 AACGTGACTCAGCGATCTCCTCAGCAGGCGAGGAGTCAATTGCTGATTTTACCGGSCAT 540
 Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180

QY 541 CGGCACACGCGGTTTCGCAACTGGAAACGGGTGCTTAATTTTCGCAATCAAACTGTGTC 600
 Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200

QY 601 CTTAAACGTGAGAAACAGGAGAAAGCTGTTTCAATTAAACGAGGATTAACCATCGAACTC 660
 Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220

QY 661 AAGTGAGCCCGACATGCTCATTAACCGTGCCTGAGCTCGCATCAATCGCTGACG 720
 Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240

QY 721 GAGATATTTCTCTGCAACACAGCACAGGAAATGCGGAGCGCGGTGGGGAAGGATCACCC 780
 Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260

QY 781 GATTACACAGTGGAAACACTGTGTAAGTTTGATTAACCGCAATTTTATTGCTACAA 840
 Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280

QY 841 CGCAGCCGACAGTTGCGCGCAGCGCCACCGCTTATTGATTTGATGTCATGGCAGCG 900
 Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300

QY 901 TTGACGCCCATCCACCGCAAAACAGCGCTATGGTGTGATGACATTAACCATCTCAGTACTG 960
 Db 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320

QY 961 TTTATTGCCGACACGATATACTTGGCAAACTCGCGCGCGCAGCTGGAGCTCAACTGG 1020
 Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340

QY 1021 ACGTTCCCGGTACGCGGATACACGCCCGCAGGTGGTGAAGTGGTGTGTTGACCGTGG 1080
 Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360


```
QY 241 GGACATTACCAACCGCAGCGCTGCTAGCCGACGATTGCTGGCGRAAAGGGCTGCGCG 300
Db 81 GlyHisTyrGlnArgGlnA-gLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGGTCCAGTCCGCGATTATTGCTGATGTCACGAGCGGTACCCGTAATAACAGGCGAA 360
Db 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCGCGCGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCCAGTCCCGATCCGTTATTATCTTAACTCTAAACTCGCGTTTCCCACTGGATAACCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AACGTGACTACGCGATCTCAGCAGGCGCAGGAGGTCAATTGCTGACTTTTACCGGCGAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAAAACGCGCTTTCGCGAATCGAAACGCGTGCTTAATTTTCGCAATCAAACTTGTC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTGAGAAACAGGACGAAGCTGTTTCAATTAACGAGCATTAACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTGAGCCCGCACATGCTCATTAACCGTGCGGTAGCCCTCGCATCAATCGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTCTCTCGCAACAGCACAGGGAATCGCGAGCGCGGTGGGGAAGGATCACCC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
QY 781 GATTACACAGTGGAAACACTTGTGTAAGTTTGATTAACCGGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
QY 841 CGCAGCCGACAGGTGCGCCGACGCGCCGACCCGTTTATGATTTGATGATGATGCGAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTGACGCCCATCCACCGCAAAACAGCGCGTATGCTGACATTCACCATTCAGTACTG 960
Db 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTGCGGACACGATATAATCTCGCAAAATCTCGCGCGCGACCTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTTCCCGTCAGCCGGATACACGCGCAGGTGTTGACTGGTGTGTTTGAACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTCGGTAAAGCATAACAGCCAGTGTGATTCAGTTTTCGTTGCTTCCAGACTTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATGCGTGATAAAACCGCGCTGTCTAATAATACGCGCGCGGAGAGGTGAACACTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGCGAGGATGTGAAGCGGAATTCGCGAGGCGATGTGTTGCTTGGCAGGTTTTCGCCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATACGACCATACCGCGCTGCGATTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
```

RESULT 6

```
Q8RKE0
ID Q8RKE0 PRELIMINARY; PRT; 432 AA.
AC Q8RKE0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90368616; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.-L.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostain K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,
RA Van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836 (1992).
DR EMBL; L03370; AAA00002.1; -.
DR HSP; P07102; 1DKM.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; HisAc phsphtse.
DR Pfam; PF00328; Acid phosphat A; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase.
SQ SEQUENCE 432 AA; 46971 MW; 5BB632D4682EABF CRC64;

Alignment Scores:
Pred. No.: 7,09e-172 Length: 432
Score: 2229.00 Matches: 429
Percent Similarity: 99.31% Conservative: 0
Best Local Similarity: 99.31% Mismatches: 3
Query Match: 94.89% Indels: 0
DB: 2 Gaps: 0

US-10-034-985-1 (1-1323) x Q8RKE0 (1-432)
QY 1 ATGAAGCGATCTTAATCCCATTTTATCTCTGATTCGTTAAACCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuLeuProGlnSerAla 20
QY 61 TTGCTCAGAGTACGCGGAGCTGAAGCTGGAAGTGTGGTGTGTCAGTCTCATGTT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerAlaHisGly 40
QY 121 GTCGCTGCTCCACCAAGGCCAGCAACTGATGAGGATCTACCCCGACGATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGCGCGTAAACCTGGGTTTCGCTGACACCGCGNGGTGGTGAAGTAACTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
QY 241 GGACATTACCAACCGCAGCGCTGCTAGCCGACGATTGCTGGCGRAAAGGGCTGCGCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGGTCCAGTCCGCGATTATTGCTGATGTCACGAGCGGTACCCGTAATAACAGGCGAA 360
Db 101 GlnSerGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCGCGCGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCCAGTCCCGATCCGTTATTATCTTAACTCTAAACTCGCGTTTCCCACTGGATAACCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AACGTGACTACGCGATCTCAGCAGGCGCAGGAGGTCAATTGCTGACTTTTACCGGCGAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAAAACGCGCTTTCGCGAATCGAAACGCGTGCTTAATTTTCGCAATCAAACTTGTC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTGAGAAACAGGACGAAGCTGTTTCAATTAACGAGCATTAACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTGAGCCCGCACATGCTCATTAACCGTGCGGTAGCCCTCGCATCAATCGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTCTCTCGCAACAGCACAGGGAATCGCGAGCGCGGTGGGGAAGGATCACCC 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
QY 781 GATTACACAGTGGAAACACTTGTGTAAGTTTGATTAACCGGCAATTTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
QY 841 CGCAGCCGACAGGTGCGCCGACGCGCCGACCCGTTTATGATTTGATGATGATGCGAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTGACGCCCATCCACCGCAAAACAGCGCGTATGCTGACATTCACCATTCAGTACTG 960
Db 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTGCGGACACGATATAATCTCGCAAAATCTCGCGCGCGACCTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTTCCCGTCAGCCGGATACACGCGCAGGTGTTGACTGGTGTGTTTGAACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTCGGTAAAGCATAACAGCCAGTGTGATTCAGTTTTCGTTGCTTCCAGACTTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATGCGTGATAAAACCGCGCTGTCTAATAATACGCGCGCGGAGAGGTGAACACTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGCGAGGATGTGAAGCGGAATTCGCGAGGCGATGTGTTGCTTGGCAGGTTTTCGCCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATACGACCATACCGCGCTGCGATTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
```

```

Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGATCCGTTATTTAATCTCTTAAACCTGGCGTTTGGCAACTGGTAACGGC 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AAGTGACTGACCGGATCTCTACAGCGGAGGAGGTCAATTCCTGACTTACCGGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGGCAACCGCGTTTCGCAACTGGAACGGGTCTTAATTTCCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTTCAATTAACGAGGATTAACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGTGAGCGCCGACATGCTCATTAACCGGTGCGTAGCCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTTCTCTGCAACAGGACAGGGAATCCGAGCGCGGTGGGGAAGGATCAC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTACACAGTGGAAACCTTGCTAAGTTTGCATTAACCGCAATTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
QY 841 CGCACGCGCAGGTTGCCCGCAGCGCCGACCCGTTATTTGATTTGATCATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
QY 901 TTGACGCGCCCATCCACGCGCAAAACAGCGGTATGGTGTGACATTACCCACTTCAGTACTG 960
Db 301 LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTCGCGACGACGATCTATCTGCGAAATCTCGCGCGCGGCTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGTTCCTCCGCTGACGCGGATAACACGCGCCGAGTGGTCAACTGGTGTGAACTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTCCGCTAAGCGATAACACCGAGTGGATTGAGTTTCGTGCTTCCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATGGCTGATAAACCGCGCTGTCTATTAAATACGCGCGCGGAGAGGTGAACCTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGSCAGGATGTGAAGAGCAAAATGCGCAGGCGATGTTGTTGGCAGGTTTACGCAA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATGAAGCACCATACCGCGGTGCGAGTTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

```

RESULT 7

```

Q8RKD9 ID Q8RKD9 PRELIMINARY; PRT; 432 AA.
AC Q8RKD9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 25, Last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.

```

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OC NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90368616; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.-L.L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appa
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase."
RL J. Bacteriol. 172:5497-5500 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostainin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,
RA Van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase."
RL J. Biol. Chem. 267:22830-22836 (1992).
DR EMBL; L03371; AAA00003.1; -
DR HSSP; P07102; 1DKO.
DR GO; GO:0003993; P:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; HisAc_phsphtse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase.
SQ SEQUENCE. 432 AA; 47033 MW; 9F29B9DF9C368175 CRC64;

```

Alignment Scores:

```

Pred. No.: 8,55e-172 Length: 432
Score: 2228.00 Matches: 429
Percent Similarity: 99.54% Conservative: 1
Best Local Similarity: 99.31% Mismatches: 2
Query Match: 94.85% Indels: 0
DB: 2 Gaps: 0

```

US-10-034-985-1 (1-1323) x Q8RKD9 (1-432)

```

QY 1 ATCAAGACGATCTAATCCCATTTTATCTTCTGATTCGTTAAACCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTCAGCGGAGCTGAAGCTGGAAGTGTGGTGAATGTCTAGTCGTCTAGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgAsnGly 40
QY 121 GTGCGTGTCTCAACCAAGCCACGCACTGATCGAGATGTCAACCCAGACGATGGCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGCGCGTAAACCTGGGTTCGTGACACCGCGNGGTGGTGAAGTAAATCGCCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
QY 241 GGACATTACCAACCGCAGCGTCTGGTAGCCGACGAGTTGTCGCGAAAAGGGCTGCCCG 300
Db 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGCTCAGTCCGGATTATTGCTGATGTCGACGAGCGTACCCGTAAACACGCGCAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCGCGCGGTGGCACCTGCTGCTGCAATAACCGTACATACCCAGGACAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCAGTCCCGATCCGTTATTTAATCTCTTAAACCTGGCGTTTGGCAACTGGTAACGGC 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AAGTGACTGACCGGATCTCTACAGCGGAGGAGGTCAATTCCTGACTTACCGGGCAT 540

```

```
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Qy 541 CGGCAACCGCGTTTCGCAACTGAAACGGGTGCTTAATTTTCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Qy 601 CTTAAACGTGAGAAACAGGACGAAGCTGTTTCATTACGAGGATTAACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Qy 661 AAGGTGAGCGCGCAATGCTCTAATTAACCGTCCGCTAAGCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaPheAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Qy 721 GAGATATTCTTCGCAACAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 780
Db 241 GluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
Qy 781 GATTACACAGTGGACACCTTGCTAAGTTTGCATTAACCGGCTCGCTAAGCTCGCATCA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
Qy 841 CGCAGCGCAGAGGTTCGCCGAGCGCGCCACCCGCTTATTGGATTGATCATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
Qy 901 TTGACGCCCATTCACCGCAAAACAGCGGTATGCTGATACATTAACCATTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyGlyValThrLeuProThrSerValLeu 320
Qy 961 TTTATTCCCGACACGATACTAATCTGGCAATCTCGCGCGGCACTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Qy 1021 ACCTTCCCGTCAGCGGATACACCGCCAGTGGTGAACCTGGTGTGAACTGGTGTGAA 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlyGluLeuValPheGluArgTrp 360
Qy 1081 CGTCGGTAAGCGATACAGCCAGTGTGATTCAGTTTCGCTGCTTCCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Qy 1141 CAGATCGGTGATAAACCGCGCTGTCTAATTAATACCGCGCGGAGAGGTGAACCTGAC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Qy 1201 CTGCGCAGATGTGAGAGCGAATTCGCGAGGATGTTGCTGTCGAGGTTTACGCA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Qy 1261 ATCGTGAATGAAGCAGCATACCGCGCTGCGAGTTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
```

RESULT 8

```
Q8RKD5 PRELIMINARY; PRT; 432 AA.
AC Q8RKD5
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90368616; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.-L.I.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
```

```
RT glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,
RA Van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836 (1992).
```

```
DR EMBL: L03375; AAA00007.1; -
DR HSP; P07102; 1DKM
DR GO: GO:0003993; F:acid phosphatase activity; IEA.
DR GO: GO:0016787; F:hydrolase activity; IEA.
DR InterPro: IPR000560; HisAc_phsphtse.
DR Pfam: PF00328; Acid_phosphat_A; 1.
DR PROSITE: PS00616; HIS_ACID_PHOSPHAT_1; 1.
KW Hydrolase.
SQ SEQUENCE 432 AA; 46990 MW; 951F3939EA9A1A47C CRC64;
```

Alignment Scores:

```
Pred. No.: 1-5e-171 Length: 432
Score: 2295.00 Matches: 429
Percent Similarity: 99.31% Conservativeness: 0
Best Local Similarity: 99.31% Mismatches: 3
Query Match: 94.72% Indels: 0
DB: 2 Gaps: 0
```

US-10-034-985-1 (1-1323) x Q8RKD5 (1-432)

```
Qy 1 ATGAAGCGATCTAATCCCATTTTATCTCTCTGATTCGTTAAACCCCGCAATCTGCA 60
Db 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuProLeuProGlnSerAla 20
Qy 61 TTCGCTCAGAGTCAGCGGAGCTGAAGCTGGAAGTGTGCTGATGTCAGTCTGCTGCT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
Qy 121 GTGCTGTCTCAACCAAGGCCAGCACTGATGAGGATGTACCCCGAGCATGGCCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Qy 181 ACCTGCGCGGTAAACCTGGGTTGGCTGACACCCGNGGTGGTGAGCTAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyLeu 80
Qy 241 GGACATTACCAACCGCAGCGTCTGCTAGCCGAGGATGCTGGCGAAAAGGGGCGCG 300
Db 81 GlyHisTyGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Qy 301 CAGTCTGCTCAGTCCGCTATTGCTGATGTGACGAGGTACCCGTAACCAAGCGGAA 360
Db 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Qy 361 GCCTTCGCGCGGCTGGCACCCTGATGTGCAATAACCGTACATACCCAGCGCATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Qy 421 TCAGTCCCGATCCGTTATTATCTTAATAAATCTGCGCTTTCGCAACTGGATAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Qy 481 AACGTGACTGACCGGATCTCTCAGCAGGAGGAGGTCAATTCGTACTTTTACCGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Qy 541 CGGCAACCGCGGTTTCGCAACTGGAACTGGAGTCTTAATTTTCGCAATCAAACTTGTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Qy 601 CTTAAACGTGAGAAACAGGACGAAGCTGTTTCATTACGAGGATTAACCATCGGAATC 660
Db 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
```

QY 661 AAGTGAGCCGACCAATGCTCATTAACCGGTGCGTAAGCCTCGCATCAATGCTGACG 720
 DB 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 QY 721 GAGATATTCTCTGTGCAACAGCAAGGAAATCCGGAGCCGGGTGGGGAAGATCACC 780
 DB 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 QY 781 GATTTCACAGTGGACACTGCTGAAGTTTCATTAACCGCGCAATTTATTTGCTACAA 840
 DB 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
 QY 841 CGACGCGCACAGGTGGCCCGCAGCCGCGCCACCCGCTTATTGATTTGATCATGGCAGCG 900
 DB 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleLysThrAla 300
 QY 901 TTGACGCCCCCATCCACCGCAAAACAGCGCTATGGTGTGACATTAACCATTCAGTACTG 960
 DB 301 LeuThrProHisProGlnLysGlnAlaTyrglyValThrLeuProThrSerValLeu 320
 QY 961 TTTATTGCCGACACGATACATACTGGCAATCTCGCGCGCAGCTGGAGCTCAACTGG 1020
 DB 321 PheIleAlaGlyAlaAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 QY 1021 ACGTCTCCCGGTGAGCGGATAACACGCGCCGAGGTGGTCAACTGGTGTGAACTGG 1080
 DB 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 QY 1081 CGTCGGTAAAGCGATAACACCGCATGAGTTCAGTTTCGTGCTTCCAGACTTACAG 1140
 DB 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 QY 1141 CAGATGCTGATAAAACCGCGCTGTCTATTAATACGCGCGCGGAGAGGTGAACTGACC 1200
 DB 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 QY 1201 CTGCGCAGGTGTGAAGAGCAAAATCGCAGCGCATGTGTTCTGTTGGCAGGTTTACGCA 1260
 DB 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 QY 1261 ATCGTGAATGAAGCAGCATACCGCGCTGCGATTG 1296
 DB 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
 RESULT 9
 Q83RW2 PRELIMINARY; PRT; 442 AA.
 AC Q83RW2
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Phosphoanhydride phosphorylase; pH 2.5 acid phosphatase;
 DE periplasmic.
 GN Name=appa; OrderedLocusNames=SF0982;
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Shigella.
 OX NCBI_TaxID=623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=301 / Serotype 2a;
 RX MEDLINE=2272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.,
 RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 RT through comparison with genomes of Escherichia coli K12 and O157.";
 RL Nucleic Acids Res. 30:4432-4441 (2002).
 DR EMBL; AE015127; AAN42610.1; -;
 DR HSSP; P07102; 1DKL.

DR GO; GO:0003993; P:acid phosphatase activity; IEA.
 DR InterPro; IPR000560; HisAc_phosphatase.
 DR Pfam; PF00328; Acid_phosphat_A; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN_1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW complete proteome.
 SQ SEQUENCE 442 AA; 48183 MW; 20BCE2C454254680 CRC64;
 Alignment Scores:
 Pred. No.: 2, 63e-171 Length: 442
 Score: 2222.00 Matches: 427
 Percent Similarity: 99.07% Conservative: 1
 Best Local Similarity: 98.84% Mismatches: 4
 Query Match: 94.59% Indels: 0
 DB: 2 Gaps: 0
 US-10-034-985-1 (1-1323) x Q83RW2 (1-442)
 QY 1 ATGAAGCGCATCTTAATCCCATTTTATCTCTTCTGATTCGTTAAACCCCGCATCTGCA 60
 DB 11 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 30
 QY 61 TTCGCTCAGAGTGAAGCGGAGCTGAAGCTGGAAGTGTGGTATTGTCAGTCTCGTCTG 120
 DB 31 PheAlaGlnSerGluProGluLeuLysLeuGluSerMetValIleValSerArgHisGly 50
 QY 121 GTGCGTGTCTCAACCAAGCCAGCCCAATCATGATCAGGATGTCAACCCAGACGATGGCCA 180
 DB 51 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 70
 QY 181 ACCTGCGCGTAAACATGGGTTGGCTGTGACACCGCGNGGTGGTGAAGTAAATCGGCTATCTC 240
 DB 71 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyLeu 90
 QY 241 GGCATTAACCAAGCCAGCGTCTGTAGCCGAGTGTGCTGCGGAAAAGGGCTGCCCG 300
 DB 91 GlyHisTyrglnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 110
 QY 301 CAGTGTGTCAGTCCGCGATTATTGTGATGTGACAGCGGTACCCCGTAAACAAACAGCGCAA 360
 DB 111 GlnSerAlaGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 130
 QY 361 GCTTCCGCGCGCGGTGGCACCTGATGTGCAATAACCGTATACCCAGCGGAGATACG 420
 DB 131 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 150
 QY 421 TCCAGTCCCGATCCGTTATTATCTCTAAACCTGGCGTTCGCAACTGGATAACGCG 480
 DB 151 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 170
 QY 481 AACGTGACTCAGCGCATCTCTCAGCAGCGGAGGAGGTCAATTTGCTGACTTTACCGGCAT 540
 DB 171 AsnValThrAspAlaIleLeuLysArgAlaGlyGlySerIleAlaAspPheThrGlyHis 190
 QY 541 CGGCAAAACCGGTTTCGCGAATCGAAGCGGTGTTAAATTTTCGCAATCAAACTTGTCG 600
 DB 191 ArgGlnThrValPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 210
 QY 601 CTTAAACGTGAGAAACAGGAGCAAGCTGTTCATTACGAGGATTAACGAGCATTAACCGAACTC 660
 DB 211 LeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 230
 QY 661 AAGTGTAGCCCGCAGCATGTCTCATTAACCGGTGCGGTAAAGCCTTCGATCAATCTGACG 720
 DB 231 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 250
 QY 721 GAGATATTCTCTGTCAACCAAGCACAGGGAATCCGAGCGCGGTGGGGAAGATCACC 780
 DB 251 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 270
 QY 781 GATTCAACAGTGGAAACACATCTGCTTAAGTTTGCATTAACCGCAATTTTATTTGCTCAA 840
 DB 271 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 290


```
QY 841 CGACGCCAGAGTTGCCCGAGCCGCCCAACCCCGTTATTGGATTTCATGGCAGCG 900
Db 291 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAsePLeuMetAlaAla 310
QY 901 TTGACGCCCATCCACCGCAAAACACGCGTGTGTGACATTACCCACTTCAGTACTG 960
Db 311 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 330
QY 961 TTTATTCCCGACACGATACCTAATCTCGCAAACTCTCGCGCGCGCACCTGGAGCTCAACTGG 1020
Db 331 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 350
QY 1021 ACGCTTCCCGGTACGCGGATACACCGCCAGGTGGTGAACCTGTTGAAACGCTGG 1080
Db 351 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 370
QY 1081 CGTCGGCTACGATACACCGCTGATTCAGGTTTCGGTCTTCAGACTTTACAG 1140
Db 371 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 390
QY 1141 CAGATGCGCTGATAAAACCGCGCTGTCTAATAATACGCGCGCGGAGAGGTGAACCTGACC 1200
Db 391 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 410
QY 1201 CTGCGCAGGATGTGAAGACGCAATTCGCGAGGCGATGTTGTTGTTGCGAGTTTACGCAA 1260
Db 411 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 430
QY 1261 ATCGTGAATCAGCACCGCATACCGCGCTGCAGTTTG 1296
Db 431 IleValAsnGluAlaArgIleProAlaCysSerLeu 442

RESULT 10
Q6RK08 PRELIMINARY; PRT; 432 AA.
ID Q6RK08 AC Q6RK08;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Acid phosphatase/phytase 2.
GN Name=appA2;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99194564; PubMed=10092520; DOI=10.1006/bbrc.1999.0361;
RA Rodriguez E., Han Y., Lei X.G.;
RT "Cloning, sequencing, and expression of an Escherichia coli acid
RT phosphatase/phytase gene (appA2) isolated from pig colon.";
RL Biochem. Biophys. Res. Commun. 257:117-123(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Lei X.G.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY496073; AAR87658.1; -
DR GO: GO:0003993; P:acid phosphatase activity; IEA.
DR InterPro: IPR005060; HisAc phosphatase.
DR Pfam: PF00328; Acid phosphatase A; 1.
DR PROSITE: PS00616; HIS ACID PHOSPHAT-2; UNKNOWN 1.
DR PROSITE: PS00778; HIS ACID PHOSPHAT-1; 1.
FT CHAIN 20 432
FT SEQUENCE 432 AA; 47041 MW; 71B0E5ED2EA2674 CRC64;

Alignment Scores:
Pred. No.: 3.16e-171 Length: 432
Score: 2221.00 Matches: 428
Percent Similarity: 99.07% Conservat: 0
Best Local Similarity: 99.07% Mismatches: 4
Query Match: 94.55% Indels: 0
DB: 2 Gaps: 0
```

```
US-10-034-985-1 (1-1323) x Q6RK08 (1-432)
QY 1 ATGAAGCGCATCTTAATCCCATTTTATCTCTTCTGATTCCGTTAACCCCGCAATCTGCA 60
Db 1 MeLysAlaIleLeuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
QY 61 TTCGCTCAGAGTGAAGCGGAGCTGAAGCTGGAAGTGTGTGATTGTGCTGCTCATGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
QY 121 GTCGCTGCTCCACCAAGCCAGCCAGCACTGATCAGAGATGTCACCCAGAGCGATGCGCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
QY 181 ACCTGCGCGGTAAACCTGGGTCGCTGACACCCCGCGGTGCTGAGCTTAATCGCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
QY 241 GGACATTTACCAACCGCAGCGCTCTGTAGCGCAGGATTGCTGCGCAAAAGGGCTGCCCG 300
Db 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGCTCAGTCCGCGATTATTGCTGATGTCGACGAGCGTACCCGTAAACAGCGCAA 360
Db 101 GlnProGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCGCGCGGTGCGCACCTGCTGCAATAACCGTACATACCCAGGCGAGATACG 420
Db 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCCAGTCCCGCATCCGTTATTATCTCTAATAAATCTGGCGTTTCCCACTGGATAACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAla 160
QY 481 AACGTGACTGACCGCATCTCTACGACGCGCAGGAGGTCAATTGCTGACTTTACGGGCAT 540
Db 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGCAACACGCGGTTTCGCAACGAGGAGCGCTCTTAATTTCCGCAATCAAACTGCTGC 600
Db 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheSerGlnLeuAsnLeuCys 200
QY 601 CTTTAAACGTGAGAAACAGGACGAAAGCTGTTCATTAAACGAGGAGCATTACCATCGGAATC 660
Db 201 LeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AGGTGAGCGCGCAGCAATGCTCATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTTCTCTGCAACAGCACAGGGAATCCGCGAGCGCGGTGGGGAAGGATCACC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetCysProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTTCACACGAGTGAACACCTTGTCTAAGTTTCATTAACGCGCAATTTTATTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
QY 841 CGACGCCAGAGTTGCCCGACCGCGCGCACCCCGCTTATGATTGATTCATGGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
QY 901 TTGACGCCCATCCACCGCAAAACACGCGTGTGTGACATTACCCACTTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTCCCGGACACGATACCTAATCTGGCAAACTCTCGCGCGCGCACCTGGAGCTCAACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGCTTCCCGGTACGCGGATACACCGCCAGGTGGTGAACCTGTTGAAACGCTGG 1080
```



```
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluValPheGluArgTrrp 360
QY 1081 CGTCCGTAAGCAGTAAACACCCAGTGGATTTCAGGTTTCGGTGGTCTTCCAGACTTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATCGCTGATAAACCGCGCTGTCTATTAATACCGCCCGGAGAGGTGAACCTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGGCAGGATCTCAGAGCCAAATCGCAGGCGATGTCGTTGCGCAGGTTTACGCA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
QY 1261 ATCGTGAATCAGCAGCATACCGCGCTGCAGTTTG 1296
Db 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432

RESULT 11
Q7UD08 PRELIMINARY; PRT; 432 AA.
AC Q7UD08;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DE Phosphoanhydride phosphorylase; pH 2.5 acid phosphatase.
GN Name=appa; OrderedLocusNames=S1048;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2457T;
RX MEDLINE=22590274; PubMed=12704152;
RX DOI=10.1128/IAI.71.5.2775-2786.2003;
RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
RA Schwartz D.C., Blattner F.R.;
RA "Complete genome sequence and comparative genomics of Shigella
RT flexneri serotype 2a strain 2457T.";
RL Infect. Immun. 71:2775-2786(2003).
DR EMBL; AE016981; AAP16495.1; -
DR HSP; P07102; 1DKL.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR InterPro; IPR000560; HisAc_phsphtse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN 1.
SQ SEQUENCE 432 AA; 47136 MW; 200442239BC66DFC CRC64;

Alignment Scores:
Pred. No.: 5.53e-171 Length: 432
Score: 2218.00 Matches: 426
Percent Similarity: 98.84% Conservative: 1
Best Local Similarity: 98.61% Mismatches: 5
Query Match: 94.42% Indels: 0
DB: 2 Gaps: 0

US-10-034-985-1 (1-1323) x Q7UD08 (1-432)
QY 1 ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCCGTTAAACCCCGCAATCGCA 60
Db 1 MetLysAlaIleuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
QY 61 TTCCTCAGATGAGCCGAGCTGAAGCTGAAGTGTGTGATTGTTCAGTCGTCTGGT 120
Db 21 PheAlaGlnSerGluProGluLeuLysLeuGluSerMetValIleValSerArgHisGly 40
QY 121 GTGGCTCTCCACCAAGGCCACCACTGATGAGGATGTCCACCCAGACGATGSCCA 180
Db 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
```

```
QY 181 ACTGCGCCGGTAAACCTGGGTTGGCTGACACCCCGNGGTGGTGAGCTAATCGCCTATCTC 240
Db 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
QY 241 GGACATTACCAACCCAGCGCTGCTAGCCGCGGATTCTGCGGAAAAAGGCTCCCG 300
Db 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
QY 301 CAGTCTGGTCAGTCGCGGATTATTGCTGATGTCGACGAGCGTACCGTAAAAACAGCGAA 360
Db 101 GlnSerAlaGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
QY 361 GCCTTCGCCCGCGGCTGTCACCTGACTGTGCAATAACCGTACATACCCAGGAGATACG 420
Db 121 AlaPheThrAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
QY 421 TCCAGTCCCGATCCGTTATTAACTCTTAAAAACTGGCGTTGCCAACTGGATACGCG 480
Db 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
QY 481 AACGTGACTGACGCGATCTCAGACGGCAGGAGGTCATTTGCTGACTTTACCGGCAT 540
Db 161 AsnValThrAspAlaIleLeuCysArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
QY 541 CGCAAAACCGCGTTTCGCAACTGGAAACCGGTGCTTAATTTCCGCAATCAAACTGTGC 600
Db 181 ArgGlnThrValPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
QY 601 CTTAAACGTGAGAAACAGCAGCAAGAACTGTTCAATTAACCGAGCATTTACCATCGGAATC 660
Db 201 LeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
QY 661 AAGGTGAGCGCCGACAAATGCTCATTAACCGGTGCGTAAAGCCTCGCATCAATGTCGACG 720
Db 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
QY 721 GAGATATTTCTCTGCAACAGCAGCAAGGAATGCCGAGCGCGGGTGGGGAAGATCACC 780
Db 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
QY 781 GATTCAACAGTGGAAACACTTGTAACTTGCATTAACCGCAATTTATTTGCTACAA 840
Db 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
QY 841 CGCACGCCAGAGGTTCCCGCAGCGCCGCGCATTCATTTGATTTGATTCATGCGCAGCG 900
Db 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
QY 901 TTGACCCCCCATCCACGCAAAACAGGCGTATGGTGTGACATACCCTTCAGTACTG 960
Db 301 LeuThrProHisProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
QY 961 TTTATTTCGGGACACGATACTAATCTGGCAAAATCTCGCGCGCGCACTGGAGCTCACTGG 1020
Db 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
QY 1021 ACGCTTCCCGGTGAGCCGGATACACGCGCCAGGTTGGTGAACCTGGTGTGTGAACGCTGG 1080
Db 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
QY 1081 CGTCCGCTAAGCGAATACACCGCATGATTCAGGTTTCGGTGGTCTTCCTCAGACTTACAG 1140
Db 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
QY 1141 CAGATCGGTGATAAACCGCGCTGTCTAATAATACCGCCCGGAGAGGTGAACCTGACC 1200
Db 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
QY 1201 CTGGCAGGATGTGAAGACGCGAAATGCGAGGCGATGTTGCTGTGGCAGGCTTTTACGCA 1260
Db 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
```

Qy	1261	ATCGTGAATGAAAGCAGCATACCGCGTGCAGTTTG	1296
Db	421	IleValAsnGluAlaArgIleProAlaCysSerLeu	432
RESULT 12			
ID	Q8CW75	PRELIMINARY;	PRT; 446 AA.
AC	Q8CW75;		
DT	01-MAR-2003	(TrEMBLrel. 23, Created)	
DT	01-MAR-2003	(TrEMBLrel. 23, Last sequence update)	
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)	
DE	Periplasmic appA protein.		
GN	Name=appA; OrderedLocuNames=c1121;		
OS	Escherichia coli O6.		
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;		
OC	Enterobacteriaceae; Escherichia.		
OX	NCBI_TaxID=217992;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=O6:H1 / CFT073 / ATCC 700928 / UPEC;		
RX	MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;		
RA	Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,		
RA	Raeko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,		
RA	Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,		
RA	Mobley H.L.T., Donnenberg M.S., Blattner F.R.;		
RT	"Extensive mosaic structure revealed by the complete genome sequence		
RT	of uropathogenic Escherichia coli.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).		
DR	EMBL; AE016758; AAM79589.1; -;		
DR	HSP; P07102; IDKQ.		
DR	GO; GO:0003993; F:acid phosphatase activity; IEA.		
DR	InterPro; IPR000560; HisAc_phspatase.		
DR	Pfam; PF00328; Acid_phosphat_A; 1.		
DR	PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.		
DR	PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.		
KW	Complete proteome.		
SQ	SEQUENCE 446 AA; 48605 MW; F1308CD691DB9F49 CRC64;		
Alignment Scores:			
Pred. No.:	1.42e-170	Length:	446
Score:	2213.00	Matches:	425
Percent Similarity:	99.07%	Conservative:	3
Best Local Similarity:	98.38%	Mismatches:	4
Query Match:	94.21%	Indels:	0
DB:	2	Gaps:	0
US-10-034-985-1 (1-1323) x Q8CW75 (1-446)			
Qy	1	ATGAAAGCATCTTAATCCCAATTTTATCTCTGATTCGGTTAAACCCGCAATCTGCA	60
Db	11	MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProLysSerAla	30
Qy	61	TTGCTCAGAGTGAGCGGAGCTGAAGCTGGAAGTGGTGGTATGTCAGTCTCATGGT	120
Db	31	PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly	50
Qy	121	GTGCGTGTCTCAACCAAGGCCACCACTGATGACGATGTACCCACAGCGATGGCCA	180
Db	51	ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro	70
Qy	181	ACCTGGCCGGTAAACATGGGTGGCTGACACCGCGNGGTGGTGAAGTAACTGCTATCTC	240
Db	71	ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyLeu	90
Qy	241	GGACATTACCAACCGCATGCTGGTAGCCGACGATTGCTGGCAAAAGGGCTGCCCG	300
Db	91	GlyHisTyrglnArgGlnArgLeuValAlaAspGlyLeuLeuThrLysLysGlyCysPro	110
Qy	301	CAGTCTGGTCAGGTCGGATTATTGCTGATGTCGACGCGGTACCCGTAACCAAGGGCAA	360
Db	111	GlnProGlyGlnValAlaIleIleSerAspValAspGluArgThrArgLysThrGlyGlu	130
Qy	361	GCCTTCGCGCGGGCTGGCACCTGACTGTGTGCAATAACCGTACATACCCAGCAGATACG	420
RESULT 13			
ID	Q7AFW3	PRELIMINARY;	PRT; 434 AA.
AC	Q7AFW3;		
DT	05-JUL-2004	(TrEMBLrel. 27, Created)	
DT	05-JUL-2004	(TrEMBLrel. 27, Last sequence update)	
DT	05-JUL-2004	(TrEMBLrel. 27, Last annotation update)	
DE	Phosphoanhydride phosphorylase.		
GN	OrderedLocuNames=ECS1136;		
OS	Escherichia coli O157:H7.		

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / RIMD 0509952 / EHEC;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayaishi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
 Kuhara S., Shiba T., Hattori M., Shinagawa H.,
 RA "Complete genome sequence of enterohemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 DR EMBL: AP002554; BAB34559.1; -.
 DR GO: GO:0003993; P:acid phosphatase activity; IEA.
 DR InterPro: IPR000560; HisAc phosphatase.
 DR Pfam: PF00328; Acid phosphatase A; 1.
 DR PROSITE: PS00616; HIS ACID PHOSPHAT 1; 1.
 DR PROSITE: PS00778; HIS ACID PHOSPHAT 2; UNKNOWN 1.
 SQ SEQUENCE 434 AA; 47337 MW; F197DF7D1869F9C4 CRC64;

Alignment Scores:
 Pred. No.: 2,46e-170 Length: 434
 Score: 2210.00 Matches: 427
 Percent Similarity: 98.39% Conservative: 0
 Best Local Similarity: 98.39% Mismatches: 5
 Query Match: 94.08% Indels: 2
 DB: 2 Gaps: 1

US-10-034-985-1 (1-1323) x Q7AFW3 (1-434)

QY 1 ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCCGTTAAACCCCGCAATCTGCA 60
 DB 1 MetLysAlaIleuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
 QY 61 TTCCTCAGAGT-----GAGCCGAGCTGAAGCTGGAAGTGTGGTATTGTGAGTCGT 114
 DB 21 PheAlaGlnSerGluProGluProGluLeuLysLeuGluSerValValIleValSerArg 40
 QY 115 CATGGTGTGGTGTCCAAACCAAGGCCACGCACTGATGAGAGTGCACCCGAGCGCA 174
 DB 41 HisGlyValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAla 60
 QY 175 TGCCCAACCTGGCGGTAAACCTGGTGGTGTGACCGGNGTGTGAGCTAATCGCC 234
 DB 61 TrpProAsnTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAla 80
 QY 235 TATCTCGGACATTAACCAACGCGCTCTGTAGCCGAGCGATTGCTGGCGAAAAGGCG 294
 DB 81 TyrLeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuThrLysLysGly 100
 QY 295 TGCCCGAGTCTGTGAGTGTGCGATTATGCTGATGTGACGAGCGTACCGGTAAACA 354
 DB 101 CysProGlnProGlyGlnValAlaIleAlaAspValAspGluArgThrArgLysThr 120
 QY 355 GGCGAAGCTTCGCGCGGCTGCGACCTGACTGTGCAATAACCGTACATACCCAGCGCA 414
 DB 121 GlyGluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAla 140
 QY 415 GATACGTCAGTCCCGATCCGTTATTTAATCTCTAAAAACTGCGGTTGCGCACTGGAT 474
 DB 141 AspThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAsp 160
 QY 475 AACCGGACGTGACTGACGCGATCCTCAGCGGCGAGGAGGTCAATGCTGACTTTACC 534
 DB 161 AsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThr 180
 QY 535 GGGCATCGGCAACCGGCTTCGCGAAGTGAACGGGTGCTTAATTTTCGCAATCAAC 594
 DB 181 GlyHisArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsn 200
 QY 595 TTGTGCCTTAAACGTGAGAAACAGGACGAAAGCTGTTTCATTAAACGCGAGCATTCACATCG 654

DB 201 LeuCysLeuAsnArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSer 220
 QY 655 GAACCTCAAGTGTAGCGCGCACAAATGCTCATTAACCGGTGCGTAAGCCTCGCATCAATG 714
 DB 221 GluLeuLysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMet 240
 QY 715 CTGACCGAGATATTCTCTCGCAACAGCACAGGGAATCCGAGCGCGGGTGGGAAGG 774
 DB 241 LeuThrGluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTyrGlyArg 260
 QY 775 ATACCGATTACACAGTGGAAACACTTGTCTAAGTTTGCATTAACCGGCAATTTATTC 834
 DB 261 IleThrAspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeu 280
 QY 835 CTACAAACGACGCGCAGAGGTTCGCCGAGCGCGCCACCCCGCTATTGATTTGATCATG 894
 DB 281 LeuGlnArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMet 300
 QY 895 GCAGCGTGTGACCGCCCATCCACCGCAAAACAGCGCTATGTTGTGACATTACCCATTCA 954
 DB 301 IleAlaLeuThrProHisProProGlnLysGlnAlaThrGlyValThrLeuProThrSer 320
 QY 955 GTACTGTTTATTCGCGACACGATACTAATCTGGCAATCTCGCGCGCGCACTGGAGCTC 1014
 DB 321 ValLeuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeu 340
 QY 1015 AACTGGACGCTTCCCGTGTGAGCGGATAACACCGCCGAGTGTGTAACCTGGTGTGAA 1074
 DB 341 AsnTrpThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGlu 360
 QY 1075 CGCTGGCGTGGTAAAGCGATTAACAGCAGTGAATTCAGGTTTCGTGTTCCAGACT 1134
 DB 361 ArgTrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThr 380
 QY 1135 TTACAGCAGATGCGTATATAAACCGCGCTGTCTATTAAATACGCGCGCGGAGGTGAA 1194
 DB 381 LeuGlnGlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLys 400
 QY 1195 CTGACCTGCGCAGGTGTGAAGAGCGAAATGCCAGCGCATGTGTTGTTGGCAGGTTTT 1254
 DB 401 LeuThrLeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPhe 420
 QY 1255 ACGCAATCTGTAATGAAGCACCGCATACCGCGTGCAGTTTG 1296
 DB 421 ThrGlnIleValAsnGluAlaArgIleProAlaCysSerLeu 434

RESULT 14
 Q8XC29 ID Q8XC29 PRELIMINARY; PRT; 444 AA.
 AC Q8XC29;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Phosphoanhydride phosphatase; pH 2.5 acid phosphatase;
 DE periplasmic.
 GN Name:appA; OrderedLocusNames=z1397;
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EHEC; ATCC 700927 / EHEC;
 RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
 RL Nature 409:529-533(2001).

```

DR EMBL; AE005292; AAG55528.1; -.
DR PIR; D85633;
DR PIR; H90770; H90770.
DR HSSP; P07102; 1DKL.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR Pfam; PF00328; Acid phosphat A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 444 AA; 48388 MW; A9AA7E6653AP247E CRC64;

Alignment Scores:
Pred. No.: 2,48e-170 Length: 444
Score: 2210.00 Matches: 427
Percent Similarity: 98.39% Conservative: 0
Best Local Similarity: 98.39% Mismatches: 5
Query Match: 94.08% Indels: 2
DB: 2 Gaps: 1

US-10-034-985-1 (1-1323) x Q8XC29 (1-444)

QY 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGGTTAAACCCGCAATCTGCA 60
DB 11 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 30
QY 61 TTCGCTCAGAGT-----GAGCCGAGCTGAAGCTGGAAGTGTGGTATGTCAGTCGT 114
DB 31 PheAlaGlnSerGluProGluProGluLeuLysLeuGluSerValValIleValSerArg 50
QY 115 CATGGTGTGGTCTCAACCAAGCCACCACTGATGAGGATGTCACCCAGCGCA 174
DB 51 HisGlyValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAla 70
QY 175 TGGCAACCTTGGCGGTAAGTGGTGTGCTGACACCGCGGTGTGAGCTAATCGCC 234
DB 71 TrpProAsnTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAla 90
QY 235 TATCTCGACATTACCAACCCAGCGTCTGGTAGCCGAGGATTCGTGGCGAAAGGCG 294
DB 91 TyrLeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuThrLysLysGly 110
QY 295 TGCCCGAGTCTGGTCAAGTTCGGATTTGCTGATGTCACGAGCGTACCCGTAACAA 354
DB 111 CysProGlnProGlyGlnValAlaIleAlaIleAlaAspValAspGluArgThrArgLysThr 130
QY 355 GCGAAGCCTTCGCGCGCGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCA 414
DB 131 GlyGluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAla 150
QY 415 GATACGTCCAGTCCCGATCCGTTATTTAACTCTTAAATACTGGCGTTTCCCACTGGAT 474
DB 151 AspThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAsp 170
QY 475 AAGCGAAGTGTACTGACGATCTCTCAGAGGCGAGGAGGTCATTTGCTGATCTTACC 534
DB 171 AsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThr 190
QY 535 GGCATCCGCAACCGGTTTCGCAACTGCAACGGGTCTTAATTTCCGCAATCAAC 594
DB 191 GlyHisArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsn 210
QY 595 TTGTGCTTAAACGTTAGAAACAGGACGAAAGCTGTTTCAATAACGAGGATTAACCATCG 654
DB 211 LeuCysLeuAsnArgGlyLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSer 230
QY 655 GAATCAAGTGTAGCCCGCAATGTCTCAATTAACCGGTGCGGTAGCCTCGCATCAATG 714
DB 231 GluLeuLysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMet 250
QY 715 CTGACCGAGATATTCTCTTCACACAGCAGGATGCGGAGCGCGGTGGGAGG 774
DB 251 LeuThrGluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArg 270

```

```

QY 775 ATCACCGATTACACACAGTGGAAACACCTTCTCTAAGTTTGTGATTAACGGCGCAATTTATTG 834
DB 271 IleThrAspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeu 290
QY 835 CTACAAACGACGCGACAGAGTTGCCCGACGCGCCACCCGTTATTGGATTGATCATG 894
DB 291 LeuGlnArgThrProGluValAlaIleArgSerArgAlaThrProLeuLeuAspLeuIleMet 310
QY 895 GCAGCGTTGACGCGCCCATCCACCGCAAAACACGCGGTATGGTGTGACATTACCCACTTCA 954
DB 311 IleAlaLeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThrSer 330
QY 955 GTACTGTTATTCGCGACACGATACTAATCTCGCAAAATCTCGCGCGCGACTGGAGCTC 1014
DB 331 ValLeuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeu 350
QY 1015 AACTGACGCGTCCCGCTCAGCGGATAACACACCGCCGAGTGTGTAACCTGGTGTGTTGAA 1074
DB 351 AsnTrpThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGlu 370
QY 1075 CGCTGCGTGGTAAAGCGATACACGCGATTCAGGTTTCGCTGGTCTTCCAGACT 1134
DB 371 ArgTrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThr 390
QY 1135 TTACACGATGCGTGATAAACCGCGCTGTCTAATAATACGCGCGCGCGAGAGGTGAAA 1194
DB 391 LeuGlnGlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLys 410
QY 1195 CTGACCTCGCAGGATGTGAAGAGCGAATACGCGAGGATGTCGTTGGCAGGTTT 1254
DB 411 LeuThrLeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPhe 430
QY 1255 ACCCAATCGTGAATGAAGCACGATACCGCGCTGCGAGTTG 1296
DB 431 ThrGlnIleValAsnGluAlaArgIleProAlaCysSerLeu 444

RESULT 15
Q676V7 PRELIMINARY; PRT; 433 AA.
ID Q676V7;
AC Q676V7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Phytase.
GN Name=PhyA;
OS Citrobacter freundii.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Citrobacter.
OX NCBI_TaxID=546;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VKPM B-4090;
RA Zinin N.V., Syneoki S.P.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY390262; AAR89622.1; -
DR InterPro; IPR000560; HisAc_phsphtse.
DR Pfam; PF00328; Acid phosphat A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
SQ SEQUENCE 433 AA; 48506 MW; A7923288F2FCFD44 CRC64;

Alignment Scores:
Pred. No.: 1,32e-102 Length: 433
Score: 1374.50 Matches: 265
Percent Similarity: 73.15% Conservative: 51
Best Local Similarity: 61.34% Mismatches: 115
Query Match: 58.51% Indels: 1
DB: 2 Gaps: 1

US-10-034-985-1 (1-1323) x Q676V7 (1-433)

QY 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTGATTCGGTTAAACCCGCAATCTGCA 60
DB 11 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 30
QY 61 TTCGCTCAGAGT-----GAGCCGAGCTGAAGCTGGAAGTGTGGTATGTCAGTCGT 114
DB 31 PheAlaGlnSerGluProGluProGluLeuLysLeuGluSerValValIleValSerArg 50
QY 115 CATGGTGTGGTCTCAACCAAGCCACCACTGATGAGGATGTCACCCAGCGCA 174
DB 51 HisGlyValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAla 70
QY 175 TGGCAACCTTGGCGGTAAGTGGTGTGCTGACACCGCGGTGTGAGCTAATCGCC 234
DB 71 TrpProAsnTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAla 90
QY 235 TATCTCGACATTACCAACCCAGCGTCTGGTAGCCGAGGATTCGTGGCGAAAGGCG 294
DB 91 TyrLeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuThrLysLysGly 110
QY 295 TGCCCGAGTCTGGTCAAGTTCGGATTTGCTGATGTCACGAGCGTACCCGTAACAA 354
DB 111 CysProGlnProGlyGlnValAlaIleAlaIleAlaAspValAspGluArgThrArgLysThr 130
QY 355 GCGAAGCCTTCGCGCGCGCTGGCACCTGACTGTGCAATAACCGTACATACCCAGGCA 414
DB 131 GlyGluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAla 150
QY 415 GATACGTCCAGTCCCGATCCGTTATTTAACTCTTAAATACTGGCGTTTCCCACTGGAT 474
DB 151 AspThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAsp 170
QY 475 AAGCGAAGTGTACTGACGATCTCTCAGAGGCGAGGAGGTCATTTGCTGATCTTACC 534
DB 171 AsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThr 190
QY 535 GGCATCCGCAACCGGTTTCGCAACTGCAACGGGTCTTAATTTCCGCAATCAAC 594
DB 191 GlyHisArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsn 210
QY 595 TTGTGCTTAAACGTTAGAAACAGGACGAAAGCTGTTTCAATAACGAGGATTAACCATCG 654
DB 211 LeuCysLeuAsnArgGlyLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSer 230
QY 655 GAATCAAGTGTAGCCCGCAATGTCTCAATTAACCGGTGCGGTAGCCTCGCATCAATG 714
DB 231 GluLeuLysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMet 250
QY 715 CTGACCGAGATATTCTCTTCACACAGCAGGATGCGGAGCGCGGTGGGAGG 774
DB 251 LeuThrGluIlePheLeuLeuGlnAlaGlnGlyMetProGluProGlyTrpGlyArg 270

```

```
Db      1 MetSerThrPheIlelleArgLeuLeuPheSerLeuLeuCysGlySerPheSerIle 20
Qy      61 TTGCTCAGAGTACGGGAGCTGAGCTGAAAGTGGTGATTTTCAGTCTCATGGT 120
Db      21 HieAlaGluGlnAsnGlyMetLysLeuGluArgValValIleValSerArgHisGly 40
Qy      121 GTGGGTCTCAACCAAGGCCAGCAACTGATGAGGATGTACCCCGACGATGGCCA 180
Db      41 ValArgAlaProThrLysPheThrProIleMetLysAspValThrProAspGlnTrpPro 60
Qy      181 ACCTGGCCGTAACATGGGTGGCTGACACCGCGNGGTGGTGAGTAAATCGCTATCTC 240
Db      61 GlnTrpAspValProLeuGlyTrpLeuThrProArgGlyGlyGluLeuValSerGluLeu 80
Qy      241 GGCATTTACCAAGCCAGCTCTGGTAGCCGAGGATGCTGGCGAAAGAGGGCTGCCCG 300
Db      81 GlyGlnTrpGlnArgLeuTrpPheThrSerLysGlyLeuLeuAsnAsnGlnThrCysPro 100
Qy      301 CAGTCTGGTTCAGGTTCGGGATTATTGCTGATGTCAGCAGCGTACCCGCTAAACAGGCGAA 360
Db      101 SerProGlyGlnValAlaValIleAlaAspThrAspGlnArgThrArgLysThrGlyGlu 120
Qy      361 GCCTTCGCCCGCGCTGGCACCTGACTGTGCAATACCGTACATACCCAGGAGATACG 420
Db      121 AlaPheLeuAlaGlyLeuAlaProLysCysGlnIleGlnValHisTyrGlnLysAspGlu 140
Qy      421 TCCAGTCCCGATCGGTATTATTAATCTCTAAACCTGGCGTTCGCCAACTGGATACGCG 480
Db      141 GluLysThrAspProLeuPheAsnProValLysMetGlyThrCysSerPheAsnThrLeu 160
Qy      481 AACGTGACTCAGCGATCCTCAGCAGGCGCAGGAGGTCAATTCGTGACTTTACCGGCGCAT 540
Db      161 LysValLysAsnAlaIleLeuGluArgAlaGlyAsnIleGluLeuTyrThrGlnArg 180
Qy      541 CGGCAAAAGCGGTTTCGCAACTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTCG 600
Db      181 TyrGlnSerSerPheArgThrLeuGluAsnValLeuAsnPheSerGlnSerGluThrCys 200
Qy      601 CTTAAACGTGAGAACAGGACGAAAGCTGTTCAATACCGCGGATTCACCATCGGAATC 660
Db      201 LysThrThrGluLysSerThrLys---CysThrLeuProGluAlaLeuProSerGluLeu 219
Qy      661 AAGGTGAGCCGCAATGCTCATTAACCGGTGCGGTAAAGCCTCGCATCAATGCTGACG 720
Db      220 LysValThrProAspAsnValSerLeuProGlyAlaTrpSerLeuSerSerThrLeuThr 239
Qy      721 GAGATATTCTCTGCAACAGACAGGAAATGCCGAGCGCGGGGTGGGGAAGATCACC 780
Db      240 GluIlePheLeuLeuGlnGluAlaGlnGlyMetProGlnValAlaTrpGlyArgIleThr 259
Qy      781 GATTACACACAGTGAACACCTTGCTAAGTTGTCATACCGCAATTTATTGCTACAA 840
Db      260 GlyGluLysGluTrpArgAspLeuLeuSerLeuHisAsnAlaGlnPheAspLeuLeuGln 279
Qy      841 CGCACCGCAGAGTTCGCCGAGCGGCCACCCCGTTATTGGATTGATCATGGCAGCG 900
Db      280 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspMetIleAspThrAla 299
Qy      901 TTGACGCCCATCCACCGCAAAAACAGCGGTATCGGTGTGACATTACCCACTTCAGTACTG 960
Db      300 LeuLeuThrAsnGlyThrThrGluAsnArgTyrGlyIleLysLeuProValSerLeuLeu 319
Qy      961 TTTATTCCCGACACGATACCTATCTGCAATCTCGCGCGCGCAGCTGGAGCTCAACTGG 1020
Db      320 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuSerGlyAlaLeuAspLeuAsnTrp 339
Qy      1021 ACGCTTCCTCCGTACGCGGATAACACGCGCGAGTGGTGAACCTGGTGTGTTTGAACGCTGG 1080
Db      340 SerLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluLysTrp 359
Qy      1081 CGTCGGGTACGGATACAGCCAGTGGATTTCAGGTTTCGGTTCGGTTCAGACTTTACAG 1140
Db      360 LysArgThrSerAspAsnThrAspTrpValGlnValSerPheValTyrGlnThrLeuArg 379
```

```
Qy      1141 CAGATCGTGATAAAAACGCCGCTGTCTATTAAATACCCCGCGAGAGGTGAAACTGACC 1200
Db      380 AspMetArgAspIleGlnProLeuSerLeuGluLysProAlaGlyLysValAspLeuLys 399
Qy      1201 CTGGCAGGATGTGAAGAGCGAAATGCGCAGGCGATGTGTTGTCGACAGTTCACGCAA 1260
Db      400 LeuIleAlaCysGluGluLysAsnSerGlnGlyMetCysSerLeuLysSerPheSerArg 419
Qy      1261 ATCGTGAATGAAGCACGACATACCGCGGTGCAGTTTG 1296
Db      420 LeuIleLysGluIleArgValProGluCysAlaVal 431
```

Search completed: March 11, 2005, 23:36:36
Job time : 277 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 22:55:05 ; Search time 168 Seconds
(without alignments)
1012.944 Million cell updates/sec

Title: US-10-034-985-2
Perfect score: 2302
Sequence: 1 MKAILPFLSLIPLTPQSA.....IVNEARIPACSLRSHHHHH 440

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2302	100.0	440	4	AAB37892
2	2302	100.0	440	5	AAE22836
3	2302	100.0	440	5	AAE15806
4	2302	100.0	440	6	ADA19446
5	2302	100.0	440	8	ADO50292
6	2282	99.1	440	8	ADO50304
7	2279	99.0	440	8	ADO50302
8	2235	97.1	432	3	AAB36257
9	2235	97.1	432	3	AAB36261
10	2235	97.1	432	3	AAB36262
11	2235	97.1	432	3	AAB36259
12	2235	97.1	432	3	AAB36258
13	2235	97.1	432	3	AAB36263
14	2235	97.1	432	3	AAB36260
15	2235	97.1	432	3	AAU77775
16	2235	97.1	432	5	AAE15807
17	2235	97.1	432	5	ABP53625
18	2235	97.1	432	6	ADA19450
19	2235	97.1	432	6	ABU15427
20	2235	97.1	432	7	ADC87745
21	2235	97.1	432	8	ADL16129
22	2235	97.1	432	8	ADO50298
23	2230	96.9	432	4	AAE02631
24	2230	96.9	432	8	ADL16131
25	2224	96.6	432	4	AAE02634

ALIGNMENTS

RESULT 1

AAAB37892
ID AAB37892 standard; protein; 440 AA.

AC AAB37892;

XX 07-MAR-2001 (first entry)

XX Escherichia coli B phytase enzyme.

DE Escherichia coli B phytase enzyme.

XX Escherichia coli B; phytase enzyme; anabolic; phytate digestion;

XX Escherichia coli B; phytase enzyme; anabolic; phytate digestion;

XX Escherichia coli.

XX WO2000071728-A1.

XX 30-NOV-2000.

XX 25-MAY-2000; 2000WO-US014846.

XX 25-MAY-1999; 99US-00318528.

XX (DIVE-) DIVERSA CORP.

XX Short JM, Kretz KA;

XX WPI; 2001-112081/12.

XX N-PSDB; AAC88885.

XX Improving the nutritional value of phytate-containing foodstuffs, using

XX phytase enzymes which catalyze the liberation of inorganic phosphate from

XX the phytates.

XX Claim 1; Fig 1; 147pp; English.

XX The present sequence is a phytase enzyme from Escherichia coli B. The

XX enzyme catalyses the liberation of inorganic phosphate from the phytate

XX in phytate-containing foodstuffs and can thus be used to improve the

XX nutritional value of phytate rich ingredients

XX Sequence 440 AA;

XX Query Match . 100.0%; Score 2302; DB 4; Length 440;

XX Best Local Similarity 100.0%; Pred.No. 2.2e-223;

XX Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESYVIVSRHGVRAPTKATQMLQDVTPDAMP 60

PI Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'donoghue E;
 XX WPI; 2002-083108/11.
 DR N-PSDB; AAD25460.
 XX
 XX New bacterial phytase for e.g. improving the nutritional value of phytate
 PT -containing foodstuffs and subsequently improving the growth performance
 PT of an organism that consumes it, or in treating animal digestive systems.
 XX
 XX Claim 1; Fig 1; 170pp; English.
 XX
 XX The patent discloses recombinant bacterial phytase from *Escherichia coli*
 CC K12 appA phytase. The enzyme has phytase activity and improved thermal
 CC tolerance when compared with wild-type phytase. It has improved protease
 CC stability at low pH. The recombinant phytase is useful for improving the
 CC nutritional value of phytate-containing foodstuffs and subsequently
 CC improving the growth performance of an organism that consumes it, in
 CC treating animal digestive systems, in feed treatment processes and for in
 CC vitro purposes related to research, discovery and development. They are
 CC also used for generating recombinant digestive system life forms, for
 CC producing or manufacturing alcoholic and non-alcoholic drinks based on
 CC the use of moulds, grains and/or plants, in biopulping and bio-bleaching
 CC where a reduction in the use of environmentally harmful chemicals that
 CC are traditionally used in the pulp and paper industry is desired and in
 CC the reduction or possible elimination of the need for mineral
 CC supplements, enzymes or therapeutic drugs for animals from the daily feed
 CC thus increasing the amount calories and nutrients present in the feed.
 CC The present sequence is *E. coli* B phytase protein
 XX Sequence 440 AA;
 SQ

Query Match 100.0%; Score 2302; DB 5; Length 440;
 Best Local Similarity 100.0%; Pred. No. 2.2e-223;
 Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAIIPLFLLIPLTQSAFAQSEPELKLVSIVSRHGVRAPTKATQLMQDVTDPWP 60
 DB 1 MKAIIPLFLLIPLTQSAFAQSEPELKLVSIVSRHGVRAPTKATQLMQDVTDPWP 60
 QY 61 TWPVKGLWLPGRGELIAYLGHYQRLVADGLLAKKGCQSGQVAVIADVDERTKTGE 120
 DB 61 TWPVKGLWLPGRGELIAYLGHYQRLVADGLLAKKGCQSGQVAVIADVDERTKTGE 120
 QY 121 AFAAGLAPDCAITVHTQADTSSPDPLENPLKTVGCQLDANVTDAILSRAGGSIAFTGH 180
 DB 121 AFAAGLAPDCAITVHTQADTSSPDPLENPLKTVGCQLDANVTDAILSRAGGSIAFTGH 180
 QY 181 ROTAFRELERVLPQSNLCIKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 DB 181 ROTAFRELERVLPQSNLCIKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 QY 241 EIFLLOAQMPGPGWGRITDSHOWNTLLSHNAQFVLLQRTPEVARSRATPLLDLIMAA 300
 DB 241 EIFLLOAQMPGPGWGRITDSHOWNTLLSHNAQFVLLQRTPEVARSRATPLLDLIMAA 300
 QY 301 LTPHPQKQYGVTLPTSVLFIAGHDNTNLNGLGALSNLTLPGQPDNTPPGGELVFERW 360
 DB 301 LTPHPQKQYGVTLPTSVLFIAGHDNTNLNGLGALSNLTLPGQPDNTPPGGELVFERW 360
 QY 361 RRLSDNSQWTVSLVFTQLOMRDKTPTLSLNTTPPGVKLFIAGCEERNAQCMCSLAGFTQ 420
 DB 361 RRLSDNSQWTVSLVFTQLOMRDKTPTLSLNTTPPGVKLFIAGCEERNAQCMCSLAGFTQ 420
 QY 421 IVNEARIPACSLRSHHHHH 440
 DB 421 IVNEARIPACSLRSHHHHH 440

RESULT 4
 ADA19446
 ID ADA19446 standard; protein; 440 AA.
 XX
 AC ADA19446;

XX 20-NOV-2003 (first entry)
 XX E. coli B phytase.
 DE
 XX Phytase; enzyme; phytate; appA gene; animal feed; inorganic phosphate;
 KW digestion enhancement; transgenic; thermal tolerance; protease stability.
 XX
 OS *Escherichia coli* B.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 72 /note= "Encoded by CGN"
 XX
 PN US2002136754-A1.
 XX
 PD 26-SEP-2002.
 XX
 XX 24-MAY-2001; 2001US-00866379.
 XX
 XX 13-AUG-1997; 97US-00910798.
 PR 01-MAR-1999; 99US-00259214.
 PR 13-APR-1999; 99US-00291931.
 PR 25-MAY-1999; 99US-00318528.
 PR 25-MAY-2000; 2000US-00580515.
 XX
 (SHOR/) SHORT J M.
 PA (KRET/) KRETZ K A.
 PA (GRAY/) GRAY K A.
 PA (BART/) BARTON N R.
 PA (GARR/) GARRETT J B.
 PA (DONO/) O' DONOGHUE E.
 PA (MATH/) MATHUR E J.
 XX
 PI Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB;
 PI O' Donoghue E, Mathur EJ;
 XX WPI; 2003-040002/03.
 XX
 XX Isolated *Escherichia coli* polynucleotide encoding a modified phytase
 PT enzyme, useful in the production of animal feed, for improving the
 PT nutritional value of phytate-containing foodstuff and for enhancing
 PT digestion in humans and animals.
 XX
 PS Claim 1; Fig 1; 62pp; English.
 XX
 CC The invention relates to an isolated *Escherichia coli* polynucleotide
 CC encoding a phytase enzyme appearing as ADA19450 and having amino acids
 CC modification from W68E, Q84W, A95P, K97C, S168E, R180Y, N226C or Y277D.
 CC Also included the *E. coli* appA gene ADA19449 (or an oligonucleotide
 CC derived from it) or its mutant sequence ADA19452, expression vectors,
 CC host cells, a method of improving nutritional value of a phytate-
 CC containing foodstuff by contacting the phytate-containing foodstuff with
 CC a pure phytase enzyme mutated phytase (where the phytase enzyme catalyzes
 CC the liberation of inorganic phosphate from the phytate in the phytate-
 CC containing foodstuff), a method to produce an animal feed containing a
 CC microbial phytase (comprising culturing the plant cell, plant part or
 CC plant under conditions where the nucleotide sequence is expressed and
 CC converting the plant cells, plants or plant into a composition for animal
 CC feed), a feed composition for animals (comprising the plant seeds, plant
 CC cells, plant parts or plants in admixture with a phytate-containing
 CC foodstuff), a method to treat a human or an animal able to benefit from
 CC digestive enhancement by the activity of an exogenous phytase enzymes
 CC comprising administering to the human or animal the plant seed, plant
 CC cells, plant parts or plants of a transgenic plant which is modified to
 CC contain an expression system which expresses a nucleotide sequence
 CC encoding a phytase enzyme, a transgenic non-human organism whose genome
 CC comprising a heterologous nucleic acid sequence encoding a polypeptide
 CC having phytase activity. The phytase enzyme is useful for improving the
 CC nutritional value of phytate-containing foodstuff, in the production of
 CC animal feed and for enhancing digestion in humans and animals. The
 CC invented method improves thermal tolerance and protease stability. It
 CC also improves the feeding value of phytate rich ingredients. The present

```
CC sequence represents E. coli B wild-type phytase.
XX
SQ Sequence 440 AA;

Query Match 100.0%; Score 2302; DB 6; Length 440;
Best Local Similarity 100.0%; Pred. No. 2.2e-223;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDWP 60
Db 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDWP 60
Qy 61 TWPVKLGLTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVQVAILIADVDRTRKTGE 120
Db 61 TWPVKLGLTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVQVAILIADVDRTRKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
Qy 181 ROTAFRELERVLPPOSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 ROTAFRELERVLPPOSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLOAQOGMPGPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFLLOAQOGMPGPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Qy 301 LTPHPKQKQAYGVTLPTSVLFIAGHDNTNLANLGALBNLWTLPGQPDNTPPGGELVFERW 360
Db 301 LTPHPKQKQAYGVTLPTSVLFIAGHDNTNLANLGALBNLWTLPGQPDNTPPGGELVFERW 360
Qy 361 RLSDNSQWITQVSLVFTQLOOMRDKTPLSLNTPPGVKLTLAGCEERNAQGMCSLAGFTQ 420
Db 361 RLSDNSQWITQVSLVFTQLOOMRDKTPLSLNTPPGVKLTLAGCEERNAQGMCSLAGFTQ 420
Qy 421 IVNEARIPACSLRSHHHHH 440
Db 421 IVNEARIPACSLRSHHHHH 440

RESULT 5
ID ADO50292 standard; protein; 440 AA.
AC ADO50292;
XX
XX 29-JUL-2004 (first entry)
XX
XX Escherichia coli B phytase enzyme.
XX
XX Phytase; bacteria; thermal tolerance; protease stability; foodstuff;
XX phytate; animal feed; fish feed; dough; baking; enzyme.
XX
XX Escherichia coli; B.
XX
XX Key Location/Qualifiers
XX Misc-difference 72 /note= "Encoded by CGN"
XX
XX US2004091968-A1.
XX
XX 13-MAY-2004.
XX
XX 20-JUN-2003; 2003US-00601319.
XX
XX 13-AUG-1997; 97US-00910798.
XX 01-MAR-1999; 99US-00259214.
XX 13-APR-1999; 99US-00291931.
XX 25-MAY-1999; 99US-00318528.
XX 25-MAY-2000; 2000US-00580515.
XX 24-MAY-2001; 2001US-00866379.
XX
XX
```

(SHORT) SHORT J M.
(KRETZ) KRETZ K.
(GRAY) GRAY K A.
(BART) BARTON N R.
(GARR) GARRETT J B.
(ODON) O'DONOGHUE E.
(MATH) MATHER E J.
Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
Mather EJ;
WPI; 2004-374952/35.
N-PSDB; ADO50291.
Producing phytase, involves providing nucleic acid derived from bacteria
encoding polypeptide having phytase activity, and expressing nucleic acid
in yeast.
Claim 3; SEQ ID NO 2; 74pp; English.
The invention relates to a method for producing phytase that involves
providing a nucleic acid encoding phytase derived from a bacteria, and
expressing the nucleic acid in a yeast under conditions that allow
expression of the enzyme in the yeast. The invention also relates to
modified phytase enzyme which has improved thermal tolerance and protease
stability at low pH. The phytase enzyme can be used in foodstuffs to
improve the feeding value of phytate rich ingredients, and in diet of
numerous animals including mammals, fowls and fishes, commercially
significant mammals such as pigs, goats, laboratory rodents, commercially
significant avian species such as chicken, ducks, doves, parrot, etc.,
commercially farmed fish such as guppy, zebrafish, molly, swordtail,
etc., in dough making and baking, in dietary aids for animals. The method
provides easy manufacture of the active ingredient loaded biocompatible
composition, higher yields and loading efficiency. The phytase
incorporated in to the dietary aid is safe for animals. The present
sequence is Escherichia coli B phytase enzyme.

Query Match 100.0%; Score 2302; DB 8; Length 440;
Best Local Similarity 100.0%; Pred. No. 2.2e-223;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDWP 60
Db 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDWP 60
Qy 61 TWPVKLGLTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVQVAILIADVDRTRKTGE 120
Db 61 TWPVKLGLTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVQVAILIADVDRTRKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
Qy 181 ROTAFRELERVLPPOSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 ROTAFRELERVLPPOSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLOAQOGMPGPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFLLOAQOGMPGPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Qy 301 LTPHPKQKQAYGVTLPTSVLFIAGHDNTNLANLGALBNLWTLPGQPDNTPPGGELVFERW 360
Db 301 LTPHPKQKQAYGVTLPTSVLFIAGHDNTNLANLGALBNLWTLPGQPDNTPPGGELVFERW 360
Qy 361 RLSDNSQWITQVSLVFTQLOOMRDKTPLSLNTPPGVKLTLAGCEERNAQGMCSLAGFTQ 420
Db 361 RLSDNSQWITQVSLVFTQLOOMRDKTPLSLNTPPGVKLTLAGCEERNAQGMCSLAGFTQ 420
Qy 421 IVNEARIPACSLRSHHHHH 440
Db 421 IVNEARIPACSLRSHHHHH 440

Db 421 INNEARIPACSLRSHHHHH 440

RESULT 6

AD050304

ID AD050304 standard; protein; 440 AA.

XX AD050304;

XX

DT 29-JUL-2004 (first entry)

XX

DE Escherichia coli W phytase 875PH2 mutant enzyme.

XX

XX Phytase; bacteria; thermal tolerance; protease stability; foodstuff;

KW phytate; animal feed; fish feed; dough; baking; mutant; muten.

XX

OS Escherichia coli.

XX

XX

XX

PH Key Location/Qualifiers

FT Misc-difference 72 /note= "Encoded by CGN"

FT Misc-difference 160 /note= "E. coli B phytase Ala replaced with Ser"

FT Misc-difference 176 /note= "E. coli B phytase Asp replaced with Gly"

FT Misc-difference 298 /note= "E. coli B phytase Met replaced with Lys"

FT Misc-difference 299 /note= "E. coli B phytase Ala replaced with Thr"

XX

US2004091968-A1.

XX

PN 13-MAY-2004.

XX

XX 20-JUN-2003; 2003US-00601319.

XX

PR 13-AUG-1997; 97US-00910798.

PR 01-MAR-1999; 99US-00259214.

PR 13-APR-1999; 99US-00291931.

PR 25-MAY-1999; 99US-00318528.

PR 25-MAY-2000; 2000US-00580515.

PR 24-MAY-2001; 2001US-00866379.

XX

(SHOR/) SHORT J M.

(KRET/) KRETZ K.

(GRAY/) GRAY K A.

(BART/) BARTON N R.

(GARR/) GARRETT J B.

(ODON/) O'DONOGHUE E.

(MATH/) MATHER E J.

XX

Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;

PI Mather EJ;

PI

XX

DR WPI; 2004-374952/35.

DR N-PSDB; AD050291, AD050303.

XX

XX Producing phytase, involves providing nucleic acid derived from bacteria

PT encoding polypeptide having phytase activity, and expressing nucleic acid

PT in yeast.

XX

PS Disclosure; Page; 74pp; English.

XX

XX The invention relates to a method for producing phytase that involves

CC providing a nucleic acid encoding phytase derived from a bacteria, and

CC expressing the nucleic acid in a yeast under conditions that allow

CC expression of the enzyme in the yeast. The invention also relates to

CC modified phytase enzyme which has improved thermal tolerance and protease

CC stability at low pH. The phytase enzyme can be used in foodstuffs to

CC improve the feeding value of phytate rich ingredients, and in diet of

CC numerous animals including mammals, fowls and fishes, commercially

CC significant mammals such as pigs, goats, laboratory rodents, commercially

CC significant avian species such as chicken, ducks, doves, parrot, etc.,

CC

CC commercially farmed fish such as guppy, zebrafish, molly, swordtail,

CC etc., in dough making and baking, in dietary aids for animals. The method

CC provides easy manufacture of the active ingredient loaded biocompatible

CC composition, higher yields and loading efficiency. The phytase

CC incorporated in to the dietary aid is safe for animals. The present

CC sequence is Escherichia coli W phytase mutant enzyme. Note: The present

CC sequence is not shown in the specification but has been derived from

CC Escherichia coli B phytase AD050292.

XX

SQ Sequence 440 AA;

XX

Query Match 99.1%; Score 2282; DB 8; Length 440;

Best Local Similarity 99.1%; Pred. No. 2.4e-221;

Matches 436; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKAILIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60

DB 1 MKAILIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60

QY 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPCQGVAIADVDERTRKTGE 120

DB 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPCQGVAIADVDERTRKTGE 120

QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTCVCLDNNVTDAILSRAGGSIAFTGH 180

DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTCVCLDNNVTDAILSRAGGSIAFTGH 180

QY 181 ROTAFRELERVLNFPQSNICLKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLT 240

DB 181 ROTAFRELERVLNFPQSNICLKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLT 240

QY 241 EIFLLOQAQMPGPGWGRITDSHWNLTLLSHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300

DB 241 EIFLLOQAQMPGPGWGRITDSHWNLTLLSHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300

QY 301 LTPHPKQKAYGVTLPTSVLFTAGHDNLANLGGALELAWTLPGQDNTPPGGELVPERW 360

DB 301 LTPHPKQKAYGVTLPTSVLFTAGHDNLANLGGALELAWTLPGQDNTPPGGELVPERW 360

QY 361 RRLSDNSQMIQSVLFTQQRMDKTPSLINTPPGVEVKLTLAGCEERNAQCMCSLAGFTQ 420

DB 361 RRLSDNSQMIQSVLFTQQRMDKTPSLINTPPGVEVKLTLAGCEERNAQCMCSLAGFTQ 420

QY 421 INNEARIPACSLRSHHHHH 440

DB 421 INNEARIPACSLRSHHHHH 440

RESULT 7

AD050302

ID AD050302 standard; protein; 440 AA.

XX AD050302;

XX

DT 29-JUL-2004 (first entry)

XX

DE Kangaroo rat Escherichia coli phytase 872PH1 mutant enzyme.

XX

XX Phytase; bacteria; thermal tolerance; protease stability; foodstuff;

KW phytate; animal feed; fish feed; dough; baking; mutant; muten.

XX

OS Escherichia coli.

XX

XX

PH Key Location/Qualifiers

FT Misc-difference 72 /note= "Encoded by CGN"

FT Misc-difference 176 /note= "E. coli B phytase Asp replaced with Gly"

FT Misc-difference 298 /note= "E. coli B phytase Met replaced with Lys"

FT Misc-difference 299 /note= "E. coli B phytase Ala replaced with Thr"

FT Misc-difference 312 /note= "E. coli B phytase Ala replaced with Thr"

/note= "E. coli B phytase Gly replaced with Ser"	
FT	US2004091968-A1.
XX	13-MAY-2004.
XX	20-JUN-2003; 2003US-00601319.
XX	13-AUG-1997; 97US-00910798.
PR	01-MAR-1999; 99US-00259214.
PR	13-APR-1999; 99US-00291931.
PR	25-MAY-1999; 99US-00318528.
PR	25-MAY-2000; 2000US-00580515.
PR	24-MAY-2001; 2001US-00866379.
XX	(SHOR/) SHORT J M.
PA	(KRET/) KRETZ K.
PA	(GRAY/) GRAY K A.
PA	(BART/) BARTON N R.
PA	(GARR/) GARRETT J B.
PA	(ODON/) O'DONOGHUE E.
PA	(MATH/) MATHER E J.
XX	Short JM, Kretz K, Gray KA, Barton NR, Garrett JB, O'donoghue E;
PI	Mather EJ;
XX	WPI; 2004-374952/35.
DR	N-PSDB; ADOS0301.
XX	Producing phytase, involves providing nucleic acid derived from bacteria
PT	encoding polypeptide having phytase activity, and expressing nucleic acid
PT	in yeast.
XX	Disclosure; Page; 74pp; English.
XX	The invention relates to a method for producing phytase that involves
CC	providing a nucleic acid encoding phytase derived from a bacteria, and
CC	expressing the nucleic acid in a yeast under conditions that allow
CC	expression of the enzyme in the yeast. The invention also relates to
CC	modified phytase enzyme which has improved thermal tolerance and protease
CC	stability at low pH. The phytase enzyme can be used in feedstuffs to
CC	improve the feeding value of phytate rich ingredients, and in diet of
CC	numerous animals including mammals, fowls and fishes, commercially
CC	significant mammals such as pigs, goats, laboratory rodents, commercially
CC	significant avian species such as chicken, ducks, doves, parrot, etc.,
CC	commercially farmed fish such as guppy, zebrafish, molly, swordtail,
CC	etc., in dough making and baking, in dietary aids for animals. The method
CC	provides easy manufacture of the active ingredient loaded biocompatible
CC	composition, higher yields and loading efficiency. The phytase
CC	incorporated in to the dietary aid is safe for animals. The present
CC	sequence is Kangaroo rat Escherichia coli phytase mutant enzyme. Note:
CC	The present sequence is not shown in the specification but has been
CC	derived from Escherichia coli B phytase sequence ADOS0292.
XX	Sequence 440 AA;
Query Match 99.08; Score 2279; DB 8; Length 440;	
Best Local Similarity 99.18; Pred. No. 4.7e-221; Mismatches 0; Gaps 0;	
Matches 436; Conservative 0; Indels 4; Indels 0; Gaps 0;	
Qy	1 MKAILIFLLSLLIPTQSAFAQSEPELKLESVVIVSRHGVRAFTKATQMLQMDVTPDAMP 60
Db	1 MKAILIFLLSLLIPTQSAFAQSEPELKLESVVIVSRHGVRAFTKATQMLQMDVTPDAMP 60
Qy	61 TWPVKLWLTFRGELLTAYLGHYORQLVADGLLAKGCPQSGQVAILADVDERTRTGE 120
Db	61 TWPVKLWLTFRGELLTAYLGHYORQLVADGLLAKGCPQSGQVAILADVDERTRTGE 120
Qy	121 AFAAGLAPDCAITVHTQADTSSPPDLENPLKTVGCQLDNANVTDAILSRAGGSITADPTGH 180
Db	121 AFAAGLAPDCAITVHTQADTSSPPDLENPLKTVGCQLDNANVTDAILSRAGGSITADPTGH 180
Qy	181 RQTAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

```

QY 1 MKAIIIPFLSLLIPTPSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
D 1 MKAIIIPFLSLLIPTPSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
QY 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVVIAIADVDERTRKTGE 120
D 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVVIAIADVDERTRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAIILSRAGGSIAFTGH 180
D 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAIILSRAGGSIAFTGH 180
QY 181 RQAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
D 181 RQAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLQQAQGMPEPGMGRIITDSHOWNTLLSHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
D 241 EIFLQQAQGMPEPGMGRIITDSHOWNTLLSHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
D 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
QY 361 RRLSDNSQWIOVSLVFOTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
D 361 RRLSDNSQWIOVSLVFOTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
D 421 IVNEARIPACSL 432

```

RESULT 9

```

AAB36261
ID AAB36261 standard; protein; 432 AA.
XX AC
AC AAB36261;
XX DT
DT 12-SEP-2003 (revised)
DT 20-FEB-2001 (first entry)
XX DE
DE R15/APPA plasmid translated sequence.
XX KW
KW Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW environmental pollution; pig.
XX OS
OS Rattus sp.
OS Escherichia coli.
OS Chimeric.
XX PN
PN WO200064247-A1.
XX PD
PD 02-NOV-2000.
XX PF
PF 20-APR-2000; 2000WO-CA000430.
XX PR
PR 23-APR-1999; 99US-0130508P.
XX PA
PA (UYGU-) UNIV GUELPH.
XX PI
PI Forsberg CW, Golovan S, Phillips JP;
XX DR
DR WPI; 2000-687245/67.
DR N-PSDB; AAC86298.
XX PT
PT Transgenic non-human animal for gastrointestinal tract specific
PT expression of a protein, preferably phytase, comprises a nucleic acid
PT sequence including a heterologous transgene construct encoding the
PT protein.
XX PS
PS Disclosure; Fig 21; 152pp; English.

```

```

XX CC The present invention provides transgenic animals which produce desired
CC proteins, in this case pigs which express phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC on 12-SEP-2003 to standardise OS field)
XX SQ
SQ Sequence 432 AA;

```

Query Match 97.1%; Score 2235; DB 3; Length 432;

Best Local Similarity 99.5%; Pred. No. 1.3e-216;

Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 MKAIIIPFLSLLIPTPSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
D 1 MKAIIIPFLSLLIPTPSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
QY 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVVIAIADVDERTRKTGE 120
D 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVVIAIADVDERTRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAIILSRAGGSIAFTGH 180
D 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAIILSRAGGSIAFTGH 180
QY 181 RQAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
D 181 RQAFRELERVLNFPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLQQAQGMPEPGMGRIITDSHOWNTLLSHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
D 241 EIFLQQAQGMPEPGMGRIITDSHOWNTLLSHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
D 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVFERW 360
QY 361 RRLSDNSQWIOVSLVFOTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
D 361 RRLSDNSQWIOVSLVFOTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
D 421 IVNEARIPACSL 432

```

RESULT 10

```

AAB36262
ID AAB36262 standard; protein; 432 AA.
XX AC
AC AAB36262;
XX DT
DT 12-SEP-2003 (revised)
DT 20-FEB-2001 (first entry)
XX DE
DE SV40/APPA plasmid translated sequence.
XX KW
KW Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW environmental pollution; pig.
XX OS
OS Simian virus 40.
OS Escherichia coli.
OS Chimeric.
XX PN
PN WO200064247-A1.
XX PD
PD 02-NOV-2000.
XX PF
PF 20-APR-2000; 2000WO-CA000430.
XX PR
PR 23-APR-1999; 99US-0130508P.

```

XX (UYGU-) UNIV GUELP.
XX Forsberg CW, Golovan S, Phillips JP;
XX WPI; 2000-687245/67.
XX N-PSDB; AAC68299.
XX Transgenic non-human animal for gastrointestinal tract specific
PT expression of a protein, preferably phytase, comprises a nucleic acid
PT sequence including a heterologous transgene construct encoding the
PT protein.
XX Disclosure; Fig 22; 152pp; English.
XX The present invention provides transgenic animals which produce desired
CC proteins, in this case pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC on 12-SEP-2003 to standardise OS field)
XX SQ Sequence 432 AA;
Query Match 97.1%; Score 2235; DB 3; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.3e-216;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MKAIIPLFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDWP 60
Db 1 MKAIIPLFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDWP 60
QY 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSGOVAILIADVDRTRKTGE 120
Db 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSGOVAILIADVDRTRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180
QY 181 RQTAFLERELVNFPPQSNCLKREKQDESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQTAFLERELVNFPPQSNCLKREKQDESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOAQMGPEPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSATPLDLLIKTA 300
Db 241 EIFLLOAQMGPEPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSATPLDLLIKTA 300
QY 301 LTPHPKQAYGVTLPTSVLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGGELVFERW 360
Db 301 LTPHPKQAYGVTLPTSVLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGGELVFERW 360
QY 361 RRLSDNSQWIOVSLVFQTLQMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
Db 361 RRLSDNSQWIOVSLVFQTLQMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
QY 421 INVNEARIPACSL 432
Db 421 INVNEARIPACSL 432
RESULT 11
AAB36259 standard; protein; 432 AA.
XX AAB36259;
XX AC
XX 12-SEP-2003 (revised)
DT 20-FEB-2001 (first entry)
XX R15/APPA plasmid translated sequence.
XX

KW Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW environmental pollution; pig.
XX Rattus sp.
OS Escherichia coli.
OS Chimeric.
XX WO200064247-A1.
XX 02-NOV-2000.
XX 20-APR-2000; 2000WO-CA000430.
XX 23-APR-1999; 99US-0130508P.
XX (UYGU-) UNIV GUELP.
XX Forsberg CW, Golovan S, Phillips JP;
XX WPI; 2000-687245/67.
XX N-PSDB; AAC68299.
XX Transgenic non-human animal for gastrointestinal tract specific
PT expression of a protein, preferably phytase, comprises a nucleic acid
PT sequence including a heterologous transgene construct encoding the
PT protein.
XX Disclosure; Fig 19; 152pp; English.
XX The present invention provides transgenic animals which produce desired
CC proteins, in this case pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC on 12-SEP-2003 to standardise OS field)
XX SQ Sequence 432 AA;
Query Match 97.1%; Score 2235; DB 3; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.3e-216;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MKAIIPLFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDWP 60
Db 1 MKAIIPLFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDWP 60
QY 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSGOVAILIADVDRTRKTGE 120
Db 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSGOVAILIADVDRTRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180
QY 181 RQTAFLERELVNFPPQSNCLKREKQDESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQTAFLERELVNFPPQSNCLKREKQDESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOAQMGPEPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSATPLDLLIKTA 300
Db 241 EIFLLOAQMGPEPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSATPLDLLIKTA 300
QY 301 LTPHPKQAYGVTLPTSVLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGGELVFERW 360
Db 301 LTPHPKQAYGVTLPTSVLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGGELVFERW 360
QY 361 RRLSDNSQWIOVSLVFQTLQMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
Db 361 RRLSDNSQWIOVSLVFQTLQMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
QY 421 INVNEARIPACSL 432
Db 421 INVNEARIPACSL 432

```
Db 421 IVNEARIPACSL 432
RESULT 12
AAB36258 standard; protein; 432 AA.
ID AAB36258 standard; protein; 432 AA.
XX AAB36258;
XX 12-SEP-2003 (revised)
DT 20-FEB-2001 (first entry)
XX R15/APPA plasmid translated sequence.
DE Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
XX environmental pollution; pig.
XX Rattus sp.
OS Escherichia coli.
OS Chimeric.
XX WO200064247-A1.
XX 02-NOV-2000.
XX 20-APR-2000; 2000WO-CA000430.
XX 23-APR-1999; 99US-0130508P.
XX (UYGU-) UNIV GUELPH.
XX Forsberg CW, Golovan S, Phillips JP;
XX WPI; 2000-687245/67.
XX N-PSDB; AAC68295.
XX Transgenic non-human animal for gastrointestinal tract specific
PT expression of a protein, preferably phytase, comprises a nucleic acid
PT sequence including a heterologous transgene construct encoding the
PT protein.
XX Disclosure; Fig 18; 152pp; English.
XX The present invention provides transgenic animals which produce desired
CC proteins, in this case pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC on 12-SEP-2003 to standardise OS field)
XX Sequence 432 AA;
Query Match 97.1%; Score 2235; DB 3; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.3e-216;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MKAILIFLSLLIPTPSAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPWP 60
Db 1 MKAILIFLSLLIPTPSAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPWP 60
Qy 61 TWPVKGLWTPRGGLIAYLGHYORQLVADGLLAKKGCPOSGOVAIIADVDERTRKTGE 120
Db 61 TWPVKGLWTPRGGLIAYLGHYORQLVADGLLAKKGCPOSGOVAIIADVDERTRKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Qy 181 QRTAFRELRLVNPQSNCLKREKQDESSCLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 QRTAFRELRLVNPQSNCLKREKQDESSCLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLQQAQMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTBEVARSRATPLLDLIMAA 300
Db 241 EIFLQQAQMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTBEVARSRATPLLDLIMAA 300
Qy 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALFNWTLPGQPDNTPPGGSELVERW 360
Db 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALFNWTLPGQPDNTPPGGSELVERW 360
Qy 361 RRLSDNSQWISQVSLVFTLQOMRDKTPSLNTPPGVEVKLTLAGCEERNAQCMCSLAGFTQ 420
Db 361 RRLSDNSQWISQVSLVFTLQOMRDKTPSLNTPPGVEVKLTLAGCEERNAQCMCSLAGFTQ 420
Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432
RESULT 13
AAB36263
ID AAB36263 standard; protein; 432 AA.
XX AAB36263;
XX 12-SEP-2003 (revised)
DT 20-FEB-2001 (first entry)
XX Lama2/APPA plasmid translated sequence.
DE Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
XX environmental pollution; pig.
XX Mus musculus.
OS Escherichia coli.
OS Chimeric.
XX WO200064247-A1.
XX 02-NOV-2000.
XX 20-APR-2000; 2000WO-CA000430.
XX 23-APR-1999; 99US-0130508P.
XX (UYGU-) UNIV GUELPH.
XX Foreberg CW, Golovan S, Phillips JP;
XX WPI; 2000-687245/67.
XX N-PSDB; AAC68300.
XX Transgenic non-human animal for gastrointestinal tract specific
PT expression of a protein, preferably phytase, comprises a nucleic acid
PT sequence including a heterologous transgene construct encoding the
PT protein.
XX Disclosure; Fig 23; 152pp; English.
XX The present invention provides transgenic animals which produce desired
CC proteins, in this case pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
CC on 12-SEP-2003 to standardise OS field)
XX Sequence 432 AA;
Query Match 97.1%; Score 2235; DB 3; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.3e-216;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MKAILIFLSLLIPTPSAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPWP 60
Db 1 MKAILIFLSLLIPTPSAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPWP 60
Qy 61 TWPVKGLWTPRGGLIAYLGHYORQLVADGLLAKKGCPOSGOVAIIADVDERTRKTGE 120
Db 61 TWPVKGLWTPRGGLIAYLGHYORQLVADGLLAKKGCPOSGOVAIIADVDERTRKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Qy 181 QRTAFRELRLVNPQSNCLKREKQDESSCLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 QRTAFRELRLVNPQSNCLKREKQDESSCLTQALPSELKVSADNVSLTGAVSLASMLT 240
```


Db 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPWP 60
 Qy 61 TWPVKLGWLTTPRGELIAYLGHYQORQLVADGLLAKKGCPCQSGQVAILIADVDERTRKTGE 120
 Db 61 TWPVKLGWLTTPRGELIAYLGHYQORQLVADGLLAKKGCPCQSGQVAILIADVDERTRKTGE 120
 Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180
 Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180
 Qy 181 QTAFARELERVLNFPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 Db 181 QTAFARELERVLNFPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 Qy 241 EIFLLOAQOGMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300
 Db 241 EIFLLOAQOGMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300
 Qy 301 LTPHPKQKQAYGVTLPTSVLFIAGHDNTNLNGLGALNLTLPQPDNTTPGGELVFERW 360
 Db 301 LTPHPKQKQAYGVTLPTSVLFIAGHDNTNLNGLGALNLTLPQPDNTTPGGELVFERW 360
 Qy 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
 Db 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
 Qy 421 IVNEARIPACSL 432
 Db 421 IVNEARIPACSL 432

RESULT 14

AAB36260 standard; protein; 432 AA.
 ID AAB36260;
 AC

XX
 AC
 XX
 DT 12-SEP-2003 (revised)
 DT 20-FEB-2001 (first entry)

XX R15/APPA plasmid translated sequence.
 DE
 XX

XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KW environmental pollution; pig.
 XX

XX Rattus sp.
 OS Escherichia coli.
 OS Chimeric.

XX W0200064247-A1.
 XX

XX 02-NOV-2000.
 XX

XX 20-APR-2000; 2000WO-CA000430.
 XX

XX 23-APR-1999; 99US-0130508P.
 XX

XX (UYGU-) UNIV GUELPH.
 XX

XX Forsberg CW, Golovan S, Phillips JP;
 PI

XX WPI; 2000-687245/67.
 DR

XX N-PSDB; AAC68297.
 DR

XX Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein.
 XX

XX Disclosure; Fig 20; 152pp; English.
 FS

XX The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which expresses phytase in the salivary
 CC

CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPA phytase coding sequence. (Updated
 CC on 12-SEP-2003 to standardise OS field)
 XX

SQ Sequence 432 AA;

Query Match 97.1%; Score 2235; DB 3; Length 432;

Best Local Similarity 99.5%; Pred. No. 1.3e-216;

Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPWP 60

Db 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPWP 60

Qy 61 TWPVKLGWLTTPRGELIAYLGHYQORQLVADGLLAKKGCPCQSGQVAILIADVDERTRKTGE 120

Db 61 TWPVKLGWLTTPRGELIAYLGHYQORQLVADGLLAKKGCPCQSGQVAILIADVDERTRKTGE 120

Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180

Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH 180

Qy 181 QTAFARELERVLNFPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Db 181 QTAFARELERVLNFPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

Qy 241 EIFLLOAQOGMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300

Db 241 EIFLLOAQOGMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300

Qy 301 LTPHPKQKQAYGVTLPTSVLFIAGHDNTNLNGLGALNLTLPQPDNTTPGGELVFERW 360

Db 301 LTPHPKQKQAYGVTLPTSVLFIAGHDNTNLNGLGALNLTLPQPDNTTPGGELVFERW 360

Qy 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Db 361 RRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

Qy 421 IVNEARIPACSL 432

Db 421 IVNEARIPACSL 432

RESULT 15

AAU77775
 ID AAU77775 standard; protein; 432 AA.
 XX

XX AAU77775;
 XX

DT 05-JUN-2002 (first entry)
 XX

XX Phytase associated protein.
 DE

XX Phytase.
 KW

XX Unidentified.
 OS

XX KR99086028-A.
 PN

XX 15-DEC-1999.
 PD

XX 25-MAY-1998; 98KR-00018810.
 PF

XX 25-MAY-1998; 98KR-00018810.
 PR

XX (WOJ-) WOJIN CO LTD.
 PA

XX Bae HD, Forceburgh CW, Goloben S, Cheng KJ;
 PI

XX WPI; 2000-645078/62.
 DR

DR N-PSDB; ABK12514.

XX	Novel phytase gene, recombinant phytase and usage thereof.
FT	Disclosure; Fig 3; 10pp; Korean.
XX	The invention relates to a novel phytase gene, a recombinant phytase gene
PS	and their uses. This is the amino acid sequence of the phytase associated
XX	protein described in the invention
CC	Sequence 432 AA;
XX	Query Match
SQ	Best Local Similarity 97.1%; Score 2235; DB 3; Length 432;
	Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy	1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
Dd	1 MKAILIPFLSLLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
Qy	61 TWPKVLGWLTPRGELIAYLGHYQRORLVADGLLAKKGCPOSGQVAIIADVDERTRKTGE 120
Dd	61 TWPKVLGWLTPRGELIAYLGHYQRORLVADGLLAKKGCPOSGQVAIIADVDERTRKTGE 120
Qy	121 AFAAGLAPDCAITVHTQTQADTSSPPLEFNPLKTGVCOLDNANVTDAILSRAGGSTADFTGH 180
Dd	121 AFAAGLAPDCAITVHTQTQADTSSPPLEFNPLKTGVCOLDNANVTDAILSRAGGSTADFTGH 180
Qy	181 QTAFRELERVINFPQSMLCKRKQDESCSLTQALPSELKVSDNVSILTGAVSLASMLT 240
Dd	181 QTAFRELERVINFPQSMLCKREKQDESCSLTQALPSELKVSDNVSILTGAVSLASMLT 240
Qy	241 EIFLLOQAQGMPGEGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Dd	241 EIFLLOQAQGMPGEGWRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIKTA 300
Qy	301 LTPHPPKOQAYGVTLPTSVLFIAGHDTNLNIAGGALELNWTLPGOPDNTPPGGELVPERW 360
Dd	301 LTPHPPKOQAYGVTLPTSVLFIAGHDTNLNIAGGALELNWTLPGOPDNTPPGGELVPERW 360
Qy	361 RRLSDNSQIWQSVLVFOTLQQMRDKTLPISLNTPPGEVKLTLAGCEERNAQGWCSLAGFTQ 420
Dd	361 RRLSDNSQIWQSVLVFOTLQQMRDKTLPISLNTPPGEVKLTLAGCEERNAQGWCSLAGFTQ 420
Qy	421 IVNEARIPACSL 432
Dd	421 IVNEARIPACSL 432

Search completed: March 11, 2005, 23:04:48
Job time : 171 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 22:59:36 ; Search time 43 Seconds
(without alignments)
763.851 Million cell updates/sec

Title: US-10-034-985-2
Perfect score: 2302
Sequence: 1 MKAILPFLSLIPLTPQSA.....IVNEARIPACLSRSHHHHH 440

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues
Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2302	100.0	440	3	US-09-259-214-2
2	2302	100.0	440	3	US-09-318-528-2
3	2302	100.0	440	3	US-09-291-931-2
4	2302	100.0	440	4	US-09-580-515-2
5	2221	96.5	433	4	US-09-540-149A-1
6	2194.5	95.3	423	2	US-08-910-798-2
7	528	22.9	421	4	US-09-489-039A-7512
8	376	16.3	522	4	US-09-489-039A-13501
9	119	5.2	439	3	US-09-044-718-3
10	119	5.2	439	4	US-10-062-848-3
11	119	5.2	440	4	US-09-684-855-106
12	119	5.2	440	4	US-09-684-855-128
13	119	5.2	440	4	US-09-684-855-151
14	119	5.2	440	4	US-09-488-265B-6
15	119	5.2	449	3	US-09-044-718-12
16	119	5.2	449	4	US-10-062-848-12
17	119	5.2	465	3	US-08-868-435-33
18	119	5.2	465	3	US-08-744-231-33
19	119	5.2	465	3	US-09-044-718-78
20	119	5.2	465	4	US-09-636-499-6
21	119	5.2	465	4	US-09-273-871A-8
22	119	5.2	465	4	US-10-083-452-8
23	119	5.2	465	4	US-09-635-504-33
24	119	5.2	465	4	US-10-062-848-78
25	119	5.1	440	4	US-09-684-855-107
26	118	5.1	440	4	US-09-684-855-129
27	118	5.1	440	4	US-09-684-855-152

28	118	5.1	440	4	US-09-488-265B-7	Sequence 7, Appli
29	117	5.1	386	1	US-08-758-213-1	Sequence 1, Appli
30	117	5.1	386	2	US-08-692-787-48	Sequence 48, Appl
31	117	5.1	386	3	US-09-097-199-48	Sequence 48, Appl
32	117	5.1	386	4	US-09-949-016-6022	Sequence 6022, Ap
33	117	5.1	515	2	US-09-146-283-2	Sequence 2, Appli
34	117	5.1	515	3	US-08-579-823A-2	Sequence 2, Appli
35	117	5.1	515	3	US-09-344-195-2	Sequence 2, Appli
36	114.5	5.0	326	4	US-09-270-767-41623	Sequence 41623, A
37	111	4.8	440	4	US-09-684-855-108	Sequence 108, App
38	111	4.8	440	4	US-09-684-855-130	Sequence 130, App
39	111	4.8	440	4	US-09-684-855-153	Sequence 153, App
40	111	4.8	440	4	US-09-488-265B-8	Sequence 8, Appli
41	111	4.8	465	3	US-09-044-718-80	Sequence 80, Appl
42	111	4.8	465	4	US-10-062-848-80	Sequence 80, Appl
43	111	4.8	467	4	US-09-684-855-165	Sequence 165, App
44	111	4.8	467	4	US-09-488-265B-33	Sequence 33, Appl
45	110	4.8	440	4	US-09-684-855-109	Sequence 109, App

ALIGNMENTS

RESULT 1
US-09-259-214-2
; Sequence 2, Application US/09259214A
; Patent No. 6110719
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: DIVER1370-1
; CURRENT APPLICATION NUMBER: US/09/259,214A
; CURRENT FILING DATE: 1999-03-01
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-259-214-2

Query Match	100.0%	Score	2302;	DB	3;	Length	440;
Best Local Similarity	100.0%;	Pred. No.	1e-233;				
Matches	440;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MKAILPFLSLIPLTPQSAFQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP	60				
Db	1	MKAILPFLSLIPLTPQSAFQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP	60				
Qy	61	TWPKVGLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVAIADVDERTKTGE	120				
Db	61	TWPKVGLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVAIADVDERTKTGE	120				
Qy	121	AFAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH	180				
Db	121	AFAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH	180				
Qy	181	QTAFRELERVLPQSNLCIKREKQDESLTQALPSELKVSADNVSLTGAVSLASMLT	240				
Db	181	QTAFRELERVLPQSNLCIKREKQDESLTQALPSELKVSADNVSLTGAVSLASMLT	240				
Qy	241	EIFLLOAQMPGPGWGRITDSHOWNTLLSIHNAQFYLLQRTPEVARSRATPLLDLMAA	300				
Db	241	EIFLLOAQMPGPGWGRITDSHOWNTLLSIHNAQFYLLQRTPEVARSRATPLLDLMAA	300				
Qy	301	LTPHPQKQAYGVTLPSTVLFIAGHDTNLNLGSALELNWTLPGQPDNTPPGSELVERW	360				
Db	301	LTPHPQKQAYGVTLPSTVLFIAGHDTNLNLGSALELNWTLPGQPDNTPPGSELVERW	360				
Qy	361	RLSDNSQWTVQSVLVPFQTLQOMRDKTPLSLNTPPGVKLTLACGEERNAQMCSLAGTQ	420				
Db	361	RLSDNSQWTVQSVLVPFQTLQOMRDKTPLSLNTPPGVKLTLACGEERNAQMCSLAGTQ	420				

```

Db 361 RRLSDNSQWISQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Qy 421 IVNEARIPACSLRSHHHHH 440
Db 421 IVNEARIPACSLRSHHHHH 440

RESULT 2
US-09-318-528-2
; Sequence 2, Application US/09318528
; Patent No. 6183740
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/318,528
; CURRENT FILING DATE: 1999-05-25
; EARLIER APPLICATION NUMBER: 09/291,931
; EARLIER FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-318-528-2

```

```

Query Match 100.0%; Score 2302; DB 3; Length 440;
Best Local Similarity 100.0%; Pred. No. 1e-233;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVVVSRHGVRAPTKATQLMQDVTPDAMP 60
Db 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVVVSRHGVRAPTKATQLMQDVTPDAMP 60
Qy 61 TWPVKLGWLTPRGGLIAYLGHYQORQLVADGLLAKKGCPCQSQGVAIADVDERTKTGE 120
Db 61 TWPVKLGWLTPRGGLIAYLGHYQORQLVADGLLAKKGCPCQSQGVAIADVDERTKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Qy 181 QTAFARELERVLPFQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 QTAFARELERVLPFQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLQQAQMPGPGWGRITDSHWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFLQQAQMPGPGWGRITDSHWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Qy 301 LTPHPQKQAYGVTLPSTVLFIAGHDTNLNLANGLALELNTLPGQPDNTPPGGELVPERW 360
Db 301 LTPHPQKQAYGVTLPSTVLFIAGHDTNLNLANGLALELNTLPGQPDNTPPGGELVPERW 360
Qy 361 RRLSDNSQWISQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQWISQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Qy 421 IVNEARIPACSLRSHHHHH 440
Db 421 IVNEARIPACSLRSHHHHH 440

```

```

RESULT 3
US-09-291-931-2
; Sequence 2, Application US/09291931A
; Patent No. 6190897
; GENERAL INFORMATION:

```

```

; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/291,931A
; CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-291-931-2

Query Match 100.0%; Score 2302; DB 3; Length 440;
Best Local Similarity 100.0%; Pred. No. 1e-233;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVVVSRHGVRAPTKATQLMQDVTPDAMP 60
Db 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVVVSRHGVRAPTKATQLMQDVTPDAMP 60
Qy 61 TWPVKLGWLTPRGGLIAYLGHYQORQLVADGLLAKKGCPCQSQGVAIADVDERTKTGE 120
Db 61 TWPVKLGWLTPRGGLIAYLGHYQORQLVADGLLAKKGCPCQSQGVAIADVDERTKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Qy 181 QTAFARELERVLPFQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 QTAFARELERVLPFQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLQQAQMPGPGWGRITDSHWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFLQQAQMPGPGWGRITDSHWNTLLSLHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Qy 301 LTPHPQKQAYGVTLPSTVLFIAGHDTNLNLANGLALELNTLPGQPDNTPPGGELVPERW 360
Db 301 LTPHPQKQAYGVTLPSTVLFIAGHDTNLNLANGLALELNTLPGQPDNTPPGGELVPERW 360
Qy 361 RRLSDNSQWISQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQWISQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Qy 421 IVNEARIPACSLRSHHHHH 440
Db 421 IVNEARIPACSLRSHHHHH 440

RESULT 4
US-09-580-515-2
; Sequence 2, Application US/09580515
; Patent No. 6720014
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/580,515
; CURRENT FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4

```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-580-515-2

Query Match
Best Local Similarity 100.0%; Score 2302; DB 4; Length 440;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKAILIFPLSLIPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTTPDAMP 60
Db 1 MKAILIFPLSLIPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTTPDAMP 60
Qy 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPOSQGVVIAIADVDERTKTGE 120
Db 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPOSQGVVIAIADVDERTKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKGTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKGTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Qy 181 ROTAFRELRLVNPQSNLCKRKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 ROTAFRELRLVNPQSNLCKRKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLQAOQMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
Db 241 EIFLLQAOQMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
Qy 301 LTPHPKQAYGVTLPTSVLFIAGHDNLANLGALELNWTLPGQPDNTPPGGELVFERW 360
Db 301 LTPHPKQAYGVTLPTSVLFIAGHDNLANLGALELNWTLPGQPDNTPPGGELVFERW 360
Qy 361 RRLSDNSQWIOVSLVFTLOQMRDTPLSLNTTPGVEVKLTLAGCEERNAQMCSLAGFTQ 420
Db 361 RRLSDNSQWIOVSLVFTLOQMRDTPLSLNTTPGVEVKLTLAGCEERNAQMCSLAGFTQ 420
Qy 421 IVNEARIPACSLRSHHHHH 440
Db 421 IVNEARIPACSLRSHHHHH 440

RESULT 5
US-09-540-149A-1
; Sequence 1, Application US/09540149A
; Patent No. 6511699
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2791
; CURRENT APPLICATION NUMBER: US/09/540.149A
; CURRENT FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 2000-03-31
; PRIOR FILING DATE: 1999-03-31
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 433
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (433)
; OTHER INFORMATION: Xaa at position 433 in this sequence is unknown
US-09-540-149A-1

Query Match
Best Local Similarity 96.5%; Score 2221; DB 4; Length 433;
Matches 428; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKAILIFPLSLIPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTTPDAMP 60
```

```
Db 1 MKAILIFPLSLIPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTTPDAMP 60
Qy 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPOSQGVVIAIADVDERTKTGE 120
Db 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPOSQGVVIAIADVDERTKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKGTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKGTGVCOLDNANVTDAILSRAGGSIAFTGH 180
Qy 181 ROTAFRELRLVNPQSNLCKRKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 ROTAFRELRLVNPQSNLCKRKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLQAOQMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
Db 241 EIFLLQAOQMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIMAA 300
Qy 301 LTPHPKQAYGVTLPTSVLFIAGHDNLANLGALELNWTLPGQPDNTPPGGELVFERW 360
Db 301 LTPHPKQAYGVTLPTSVLFIAGHDNLANLGALELNWTLPGQPDNTPPGGELVFERW 360
Qy 361 RRLSDNSQWIOVSLVFTLOQMRDTPLSLNTTPGVEVKLTLAGCEERNAQMCSLAGFTQ 420
Db 361 RRLSDNSQWIOVSLVFTLOQMRDTPLSLNTTPGVEVKLTLAGCEERNAQMCSLAGFTQ 420
Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 6
US-08-910-798-2
; Sequence 2, Application US/08910798
; Patent No. 5876997
; GENERAL INFORMATION:
; APPLICANT: KRETZ
; TITLE OF INVENTION: NOVEL PHYTASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: California
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/910.798
; FILING DATE: August 13, 1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HAILE, PH.D., LISA A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/029001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-910-798-2

Query Match 95.3%; Score 2194.5; DB 2; Length 423;
Best Local Similarity 96.1%; Pred. No. 2.1e-222;
```



```
US-09-044-718-3
; Sequence 3, Application US/09044718
; Patent No. 6391605
; GENERAL INFORMATION:
; APPLICANT: KOSTREWA, Dirk
; APPLICANT: PASAMONTES, Luis
; APPLICANT: TOMSCHY, Andrea
; APPLICANT: van LOON, Adolphus
; APPLICANT: VOGEL, Kurt
; APPLICANT: WYSS, Markus
; TITLE OF INVENTION: MODIFIED PHYTASES
; FILE REFERENCE: Modified Phytases
; CURRENT APPLICATION NUMBER: US/09/044,718
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: EP 97810175.6
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-09-044-718-3

Query Match      5.2%; Score 119; DB 3; Length 439;
Best Local Similarity 21.7%; Pred. No. 0.0017;
Matches 102; Conservative 70; Mismatches 183; Indels 114; Gaps 24;

Qy 7 PFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPT-----KATQLMODVTPDA--- 58
Db 26 PFFSLEDELSVSSKLPK---DCRITLVQVLSRHGARYPTSSKSKYKLVTAIQANATDF 82
Qy 59 -----WPTWPKLGM--LTPRGGLIAYLG--HYQORLVADGLLAKKGCPOSQGVAAI 108
Db 83 KGKFAFLKTYNYTLGADDLTPFGQQVNSGKIFQYQYKALARSVP-----FIR 132
Qy 109 ADVDERTRKTGEAPAAGL-----APDCAITVHTQADTSPPDLNPLKTV 154
Db 133 ASGSDRVIASGEKFIQFQQAADPCATNRAAPASVII-PESETFN-----NTLDHGV 186
Qy 155 COLDNANVTDAILSRAGSSI-ADPTGHRQTAFARELERVLNFPQSNCLCKEKQDECSLT 213
Db 187 C-----TKFEASQLGDEVAANFT-----ALFAP--DIRARAEKHLPGVTLT 225
Qy 214 QALPSEL--KVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPCWGRIITDISHQWNTLLS 271
Db 226 DEDVSLMDMCSFDTVARTSDASQLSPFCOLFTHNE-----W-----KKYNYLSL 271
Qy 272 HNAQFYLLQRTPEVARSRATPLDLIMAALTTPHPQKQ-----AYGVTLPSTVLFIA 323
Db 272 --GKYGYGAGNPLGPAQGIQFTNELIARLTRSPVQDHTSTNSTLVSNPATFPLNATMYV 329
Qy 324 --GHDNTNLANLGGALELNLWTLPGQDNTPPGGLVFERWRLSD--NSQWI---QVSLVFQ 377
Db 330 DFSDHNSMWSIFFAL-----GLYNGTEPLSRSTSVEKELDCYSASVWVPFGARAYFE 382
Qy 378 TLQMRDKTPLSLNTPPGVEVKLTLAGCEERNAQGMCSLAGFTQIVNEAR 426
Db 383 TMQCKSEKPL-VRALINDRVVPLHGCDV-DKLGRCCKLNDVFKGLSWAR 429

RESULT 10
US-10-062-848-3
; Sequence 3, Application US/10062848
; Patent No. 6734004
; GENERAL INFORMATION:
; APPLICANT: KOSTREWA, Dirk
; APPLICANT: PASAMONTES, Luis
; APPLICANT: TOMSCHY, Andrea
; APPLICANT: van LOON, Adolphus
; APPLICANT: VOGEL, Kurt
; APPLICANT: WYSS, Markus
; TITLE OF INVENTION: MODIFIED PHYTASES

US-09-044-718-3
; Sequence 3, Application US/09044718
; Patent No. 6391605
; GENERAL INFORMATION:
; APPLICANT: KOSTREWA, Dirk
; APPLICANT: PASAMONTES, Luis
; APPLICANT: TOMSCHY, Andrea
; APPLICANT: van LOON, Adolphus
; APPLICANT: VOGEL, Kurt
; APPLICANT: WYSS, Markus
; TITLE OF INVENTION: MODIFIED PHYTASES
; FILE REFERENCE: Modified Phytases
; CURRENT APPLICATION NUMBER: US/09/044,718
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: EP 97810175.6
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-062-848-3

Query Match      5.2%; Score 119; DB 4; Length 439;
Best Local Similarity 21.7%; Pred. No. 0.0017;
Matches 102; Conservative 70; Mismatches 183; Indels 114; Gaps 24;

Qy 7 PFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPT-----KATQLMODVTPDA--- 58
Db 26 PFFSLEDELSVSSKLPK---DCRITLVQVLSRHGARYPTSSKSKYKLVTAIQANATDF 82
Qy 59 -----WPTWPKLGM--LTPRGGLIAYLG--HYQORLVADGLLAKKGCPOSQGVAAI 108
Db 83 KGKFAFLKTYNYTLGADDLTPFGQQVNSGKIFQYQYKALARSVP-----FIR 132
Qy 109 ADVDERTRKTGEAPAAGL-----APDCAITVHTQADTSPPDLNPLKTV 154
Db 133 ASGSDRVIASGEKFIQFQQAADPCATNRAAPASVII-PESETFN-----NTLDHGV 186
Qy 155 COLDNANVTDAILSRAGSSI-ADPTGHRQTAFARELERVLNFPQSNCLCKEKQDECSLT 213
Db 187 C-----TKFEASQLGDEVAANFT-----ALFAP--DIRARAEKHLPGVTLT 225
Qy 214 QALPSEL--KVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPCWGRIITDISHQWNTLLS 271
Db 226 DEDVSLMDMCSFDTVARTSDASQLSPFCOLFTHNE-----W-----KKYNYLSL 271
Qy 272 HNAQFYLLQRTPEVARSRATPLDLIMAALTTPHPQKQ-----AYGVTLPSTVLFIA 323
Db 272 --GKYGYGAGNPLGPAQGIQFTNELIARLTRSPVQDHTSTNSTLVSNPATFPLNATMYV 329
Qy 324 --GHDNTNLANLGGALELNLWTLPGQDNTPPGGLVFERWRLSD--NSQWI---QVSLVFQ 377
Db 330 DFSDHNSMWSIFFAL-----GLYNGTEPLSRSTSVEKELDCYSASVWVPFGARAYFE 382
Qy 378 TLQMRDKTPLSLNTPPGVEVKLTLAGCEERNAQGMCSLAGFTQIVNEAR 426
Db 383 TMQCKSEKPL-VRALINDRVVPLHGCDV-DKLGRCCKLNDVFKGLSWAR 429

RESULT 11
US-09-684-855-106
; Sequence 106, Application US/09684855
; Patent No. 6599735
; GENERAL INFORMATION:
; APPLICANT: F. Hoffmann-La Roche AG
; TITLE OF INVENTION: CONTINUOUS FERMENTATION PROCESS
; FILE REFERENCE: C38435/111692
; CURRENT APPLICATION NUMBER: US/09/684,855
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: EP 00121663.9
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: EP 99120289.6
; PRIOR FILING DATE: 1999-10-11
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 106
; LENGTH: 440
; TYPE: PRT
; ORGANISM: A. fumigatus 13073
US-09-684-855-106
```

Query Match	5.2%; Score 119; DB 4; Length 440;	
Best Local Similarity	21.7%; Pred. No. 0.0017;	
Matches 102; Conservative	70; Mismatches 183; Indels 114; Gaps 24;	
QY	7 PFLSLIPLTPOSFAQSEPELKESVIVSRHGVRAPT-----KATOLMODVTPDA--- 58	
DB	27 PFFSLEDELSVSSKLPK---DCRITLVQLSRHGARYPTSSKSKYKLVTAIQANATDF 83	
QY	59 -----WPTWPVKLGW--LTPRGGLIAYLG--HYQORLVADGLLAKKGCPCQSQVAII 108	
DB	84 KGKFAFLKTYNTLTGADDLTPFGEQQLVNSGKIFQRYKALARSVVP-----FIR 133	
QY	109 ADVDERTRKTGEAFAAGL-----APCAITVHTQADTSSPDLFNPPLKTGV 154	
DB	134 ASGSDRVIASGEKFTIEGFQQAOKLADPGATNRAAPAIISVII--PESETFN-----NTLDHGV 187	
QY	155 COLDNANVTDAILSRAGSI--ADFTGHRQTAFRELVRNFPQSNCLCKREKQDESCSLT 213	
DB	188 C-----TKFEASQLGDVEAANFT-----ALFAP--DIRARAETHLPVGTILT 226	
QY	214 QALPSEL--KVSADNVSLTGAVSLASMLTEIFLLOQAQGMPEPGWGRITDHSQWNTLLSL 271	
DB	227 PFFSLEDELSVSSKLPK---DCRITLVQLSRHGARYPTSSKSKYKLVTAIQANATDF 83	
QY	59 -----WPTWPVKLGW--LTPRGGLIAYLG--HYQORLVADGLLAKKGCPCQSQVAII 108	
DB	84 KGKFAFLKTYNTLTGADDLTPFGEQQLVNSGKIFQRYKALARSVVP-----FIR 133	
QY	109 ADVDERTRKTGEAFAAGL-----APCAITVHTQADTSSPDLFNPPLKTGV 154	
DB	134 ASGSDRVIASGEKFTIEGFQQAOKLADPGATNRAAPAIISVII--PESETFN-----NTLDHGV 187	
QY	155 COLDNANVTDAILSRAGSI--ADFTGHRQTAFRELVRNFPQSNCLCKREKQDESCSLT 213	
DB	188 C-----TKFEASQLGDVEAANFT-----ALFAP--DIRARAETHLPVGTILT 226	
QY	214 QALPSEL--KVSADNVSLTGAVSLASMLTEIFLLOQAQGMPEPGWGRITDHSQWNTLLSL 271	
DB	227 PFFSLEDELSVSSKLPK---DCRITLVQLSRHGARYPTSSKSKYKLVTAIQANATDF 83	
QY	59 -----WPTWPVKLGW--LTPRGGLIAYLG--HYQORLVADGLLAKKGCPCQSQVAII 108	
DB	84 KGKFAFLKTYNTLTGADDLTPFGEQQLVNSGKIFQRYKALARSVVP-----FIR 133	
QY	109 ADVDERTRKTGEAFAAGL-----APCAITVHTQADTSSPDLFNPPLKTGV 154	
DB	134 ASGSDRVIASGEKFTIEGFQQAOKLADPGATNRAAPAIISVII--PESETFN-----NTLDHGV 187	
QY	155 COLDNANVTDAILSRAGSI--ADFTGHRQTAFRELVRNFPQSNCLCKREKQDESCSLT 213	
DB	188 C-----TKFEASQLGDVEAANFT-----ALFAP--DIRARAETHLPVGTILT 226	
QY	214 QALPSEL--KVSADNVSLTGAVSLASMLTEIFLLOQAQGMPEPGWGRITDHSQWNTLLSL 271	
DB	227 PFFSLEDELSVSSKLPK---DCRITLVQLSRHGARYPTSSKSKYKLVTAIQANATDF 83	
QY	59 -----WPTWPVKLGW--LTPRGGLIAYLG--HYQORLVADGLLAKKGCPCQSQVAII 108	
DB	84 KGKFAFLKTYNTLTGADDLTPFGEQQLVNSGKIFQRYKALARSVVP-----FIR 133	
QY	109 ADVDERTRKTGEAFAAGL-----APCAITVHTQADTSSPDLFNPPLKTGV 154	
DB	134 ASGSDRVIASGEKFTIEGFQQAOKLADPGATNRAAPAIISVII--PESETFN-----NTLDHGV 187	
QY	155 COLDNANVTDAILSRAGSI--ADFTGHRQTAFRELVRNFPQSNCLCKREKQDESCSLT 213	
DB	188 C-----TKFEASQLGDVEAANFT-----ALFAP--DIRARAETHLPVGTILT 226	
QY	214 QALPSEL--KVSADNVSLTGAVSLASMLTEIFLLOQAQGMPEPGWGRITDHSQWNTLLSL 271	
DB	227 PFFSLEDELSVSSKLPK---DCRITLVQLSRHGARYPTSSKSKYKLVTAIQANATDF 83	
QY	59 -----WPTWPVKLGW--LTPRGGLIAYLG--HYQORLVADGLLAKKGCPCQSQVAII 108	
DB	84 KGKFAFLKTYNTLTGADDLTPFGEQQLVNSGKIFQRYKALARSVVP-----FIR 133	
QY	109 ADVDERTRKTGEAFAAGL-----APCAITVHTQADTSSPDLFNPPLKTGV 154	
DB	134 ASGSDRVIASGEKFTIEGFQQAOKLADPGATNRAAPAIISVII--PESETFN-----NTLDHGV 187	
QY	155 COLDNANVTDAILSRAGSI--ADFTGHRQTAFRELVRNFPQSNCLCKREKQDESCSLT 213	
DB	188 C-----TKFEASQLGDVEAANFT-----ALFAP--DIRARAETHLPVGTILT 226	
QY	214 QALPSEL--KVSADNVSLTGAVSLASMLTEIFLLOQAQGMPEPGWGRITDHSQWNTLLSL 271	
DB	227 PFFSLEDELSVSSKLPK---DCRITLVQLSRHGARYPTSSKSKYKLVTAIQANATDF 83	
QY	59 -----WPTWPVKLGW--LTPRGGLIAYLG--HYQORLVADGLLAKKGCPCQSQVAII 108	
DB	84 KGKFAFLKTYNTLTGADDLTPFGEQQLVNSGKIFQRYKALARSVVP-----FIR 133	
QY	109 ADVDERTRKTGEAFAAGL-----APCAITVHTQADTSSPDLFNPPLKTGV 154	
DB	134 ASGSDRVIASGEKFTIEGFQQAOKLADPGATNRAAPAIISVII--PESETFN-----NTLDHGV 187	
QY	155 COLDNANVTDAILSRAGSI--ADFTGHRQTAFRELVRNFPQSNCLCKREKQDESCSLT 213	
DB	188 C-----TKFEASQLGDVEAANFT-----ALFAP--DIRARAETHLPVGTILT 226	
QY	214 QALPSEL--KVSADNVSLTGAVSLASMLTEIFLLOQAQGMPEPGWGRITDHSQWNTLLSL 271	
DB	227 PFFSLEDELSVSSKLPK---DCRITLVQLSRHGARYPTSSKSKYKLVTAIQANATDF 83	
QY	59 -----WPTWPVKLGW--LTPRGGLIAYLG--HYQORLVADGLLAKKGCPCQSQVAII 108	
DB	84 KGKFAFLKTYNTLTGADDLTPFGEQQLVNSGKIFQRYKALARSVVP-----FIR 133	
QY	109 ADVDERTRKTGEAFAAGL-----APCAITVHTQADTSSPDLFNPPLKTGV 154	
DB	134 ASGSDRVIASGEKFTIEGFQQAOKLADPGATNRAAPAIISVII--PESETFN-----NTLDHGV 187	
QY	155 COLDNANVTDAILSRAGSI--ADFTGHRQTAFRELVRNFPQSNCLCKREKQDESCSLT 213	
DB	188 C-----TKFEASQLGDVEAANFT-----ALFAP--DIRARAETHLPVGTILT 226	
QY	214 QALPSEL--KVSADNVSLTGAVSLASMLTEIFLLOQAQGMPEPGWGRITDHSQWNTLLSL 271	
DB	227 PFFSLEDELSVSSKLPK---DCRITLVQLSRHGARYPTSSKSKYKLVTAIQANATDF 83	
QY	59 -----WPTWPVKLGW--LTPRGGLIAYLG--HYQORLVADGLLAKKGCPCQSQVAII 108	
DB	84 KGKFAFLKTYNTLTGADDLTPFGEQQLVNSGKIFQRYKALARSVVP-----FIR 133	
QY	109 ADVDERTRKTGEAFAAGL-----APCAITVHTQADTSSPDLFNPPLKTGV 154	
DB	134 ASGSDRVIASGEKFTIEGFQQAOKLADPGATNRAAPAIISVII--PESETFN-----NTLDHGV 187	
QY	155 COLDNANVTDAILSRAGSI--ADFTGHRQTAFRELVRNFPQSNCLCKREKQDESCSLT 213	
DB	188 C-----TKFEASQLGDVEAANFT-----ALFAP--DIRARAETHLPVGTILT 226	
QY	214 QALPSEL--KVSADNVSLTGAVSLASMLTEIFLLOQAQGMPEPGWGRITDHSQWNTLLSL 271	
DB	227 PFFSLEDELSVSSKLPK---DCRITLVQLSRHGARYPTSSKSKYKLVTAIQANATDF 83	
QY	59 -----WPTWPVKLGW--LTPRGGLIAYLG--HYQORLVADGLLAKKGCPCQSQVAII 108	
DB	84 KGKFAFLKTYNTLTGADDLTPFGEQQLVNSGKIFQRYKALARSVVP-----FIR 133	

Qy 324 --GHDNTLANLGGALELNTWLPQPDNTPPGGELVFERWRLSD--NSQWI---QVSLVFQ 377
Db 331 DFDHNSMVSIFPAL-----GLYNGTEPLSRSTVESAKELDGYASWVVPFGARAYFE 383
Qy 378 TLQOMRDKTPLSLNTPGGEVKLTLAGCEERNAQCMCSLAGFTQIVNEAR 426
Db 384 TMOCKSEKEPL-VRALINDRVVPLHGCDV-DKLGRCCKLNDVFKGLSWAR 430

RESULT 14

US-09-488-265B-6
; Sequence 6, Application US/09488265B
; Patent No. 6720174
; GENERAL INFORMATION:
; APPLICANT: Lehmann, Martin
; TITLE OF INVENTION: Improved Phytases
; FILE REFERENCE: 5808.200-US
; CURRENT APPLICATION NUMBER: US/09/488,265B
; CURRENT FILING DATE: 2000-01-20
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 6
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus 13073
US-09-488-265B-6

Query Match 5.2%; Score 119; DB 4; Length 440;
Best Local Similarity 21.7%; Pred. No. 0.0017;
Matches 102; Conservative 70; Mismatches 183; Indels 114; Gaps 24;
Qy 7 PFLSLILPLTPQSAFAQSEBELKLESVIVSRHGVRAPT-----KATQLMODVTPDA--- 58
Db 27 PFFSLEDELSVSSKLPK---DCRITLVQVLSRHGARYPTSSKKYKLVTAIQANATDF 83
Qy 59 -----WPTWPKLGM--LTPRGELIAYLG--HYQRQLVADGLLAKKGCPOSGQVAII 108
Db 84 KGKFAFLKTYNTLTGADDLTPFGGQQLVNSGKIFYQRYKALARSVP-----FIR 133
Qy 109 ADVDERTRKTGEAFAAGL-----APDCAITVHTQADTSSPDPLENPLKTGV 154
Db 134 ASGSDRVIASGEKFIQEQAKLADPGATNRAAPAIISVII-PESETFN-----NTLDHGV 187
Qy 155 COLDNANVTDAIILSRAGSI-ADFTGHRQTAFRELERVLPNPPQSNLCCKREKODESCSLT 213
Db 188 C-----TKFEASQLGDEVAANFT-----ALFAP--DIRARAEKHLPGVTLT 226
Qy 214 QALPSEL--KVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPCWGRITDSHOWNTLLSL 271
Db 227 DEDVVSMDMCSFDTVARTSDASQLSPFCQLFTHNE-----W-----KKYNYLQSL 272
Qy 272 HNAQFYLLQRTPEVARSRATPLDLIMAALTTPHPQKQ-----AYGVTLPSTVLFIA 323
Db 273 --GKYGYGAGNPLGPAQGGIGFTNELIARLTSRSPVQDHTSTNSTLVSNPATFPLNATMYV 330
Qy 324 --GHDNTLANLGGALELNTWLPQPDNTPPGGELVFERWRLSD--NSQWI---QVSLVFQ 377
Db 331 DFDHNSMVSIFPAL-----GLYNGTEPLSRSTVESAKELDGYASWVVPFGARAYFE 383
Qy 378 TLQOMRDKTPLSLNTPGGEVKLTLAGCEERNAQCMCSLAGFTQIVNEAR 426
Db 384 TMOCKSEKEPL-VRALINDRVVPLHGCDV-DKLGRCCKLNDVFKGLSWAR 430

RESULT 15

US-09-044-718-12
; Sequence 12, Application US/09044718
; Patent No. 6391605
; GENERAL INFORMATION:
; APPLICANT: KOSTREWA, Dirk
; PASAMONTES, Luis

; APPLICANT: TOMSCHY, Andrea
; APPLICANT: van LOON, Adolphus
; APPLICANT: VOGEL, Kurt
; APPLICANT: WYSS, Markus
; TITLE OF INVENTION: MODIFIED PHYTASES
; FILE REFERENCE: Modified Phytases
; CURRENT APPLICATION NUMBER: US/09/044,718
; CURRENT FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: EP 97810175.6
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-09-044-718-12

Query Match 5.2%; Score 119; DB 3; Length 449;
Best Local Similarity 21.7%; Pred. No. 0.0017;
Matches 102; Conservative 70; Mismatches 183; Indels 114; Gaps 24;
Qy 7 PFLSLILPLTPQSAFAQSEBELKLESVIVSRHGVRAPT-----KATQLMODVTPDA--- 58
Db 36 PFFSLEDELSVSSKLPK---DCRITLVQVLSRHGARYPTSSKKYKLVTAIQANATDF 92
Qy 59 -----WPTWPKLGM--LTPRGELIAYLG--HYQRQLVADGLLAKKGCPOSGQVAII 108
Db 93 KGKFAFLKTYNTLTGADDLTPFGGQQLVNSGKIFYQRYKALARSVP-----FIR 142
Qy 109 ADVDERTRKTGEAFAAGL-----APDCAITVHTQADTSSPDPLENPLKTGV 154
Db 143 ASGSDRVIASGEKFIQEQAKLADPGATNRAAPAIISVII-PESETFN-----NTLDHGV 196
Qy 155 COLDNANVTDAIILSRAGSI-ADFTGHRQTAFRELERVLPNPPQSNLCCKREKODESCSLT 213
Db 197 C-----TKFEASQLGDEVAANFT-----ALFAP--DIRARAEKHLPGVTLT 235
Qy 214 QALPSEL--KVSADNVSLTGAVSLASMLTEIFLLQQAQGMPEPCWGRITDSHOWNTLLSL 271
Db 236 DEDVVSMDMCSFDTVARTSDASQLSPFCQLFTHNE-----W-----KKYNYLQSL 281
Qy 272 HNAQFYLLQRTPEVARSRATPLDLIMAALTTPHPQKQ-----AYGVTLPSTVLFIA 323
Db 282 --GKYGYGAGNPLGPAQGGIGFTNELIARLTSRSPVQDHTSTNSTLVSNPATFPLNATMYV 339
Qy 324 --GHDNTLANLGGALELNTWLPQPDNTPPGGELVFERWRLSD--NSQWI---QVSLVFQ 377
Db 340 DFDHNSMVSIFPAL-----GLYNGTEPLSRSTVESAKELDGYASWVVPFGARAYFE 392
Qy 378 TLQOMRDKTPLSLNTPGGEVKLTLAGCEERNAQCMCSLAGFTQIVNEAR 426
Db 393 TMOCKSEKEPL-VRALINDRVVPLHGCDV-DKLGRCCKLNDVFKGLSWAR 439

Search completed: March 11, 2005, 23:09:24
Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 23:04:56 ; Search time 142 Seconds
(without alignments)
1022.054 Million cell updates/sec

Title: US-10-034-985-2

Perfect score: 2302

Sequence: 1 MKAILPFLSLIPLTPQSA.....IVNEARIPACSLSHHHHH 440

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1396920 seqs, 329844858 residues

Total number of hits satisfying chosen parameters: 1396920

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2302	100.0	440	9	US-09-777-566A-2
2	2302	100.0	440	9	US-09-866-379-2
3	2302	100.0	440	14	US-10-034-985-2
4	2302	100.0	440	15	US-10-430-356-2
5	2302	100.0	440	15	US-10-601-319-2
6	2335	97.1	432	9	US-09-866-379-8
7	2235	97.1	432	14	US-10-156-660-4
8	2235	97.1	432	15	US-10-282-122A-43351
9	2235	97.1	432	15	US-10-601-319-8
10	2235	97.1	432	16	US-10-472-317-41
11	2221	96.5	432	15	US-10-284-962-3
12	2221	96.5	432	15	US-10-284-962-14
13	2221	96.5	433	14	US-10-266-041-1
					Sequence 2, Appli
					Sequence 2, Appli
					Sequence 2, Appli
					Sequence 2, Appli
					Sequence 8, Appli
					Sequence 4, Appli
					Sequence 43351, A
					Sequence 8, Appli
					Sequence 41, Appli
					Sequence 3, Appli
					Sequence 14, Appli
					Sequence 1, Appli

14	2197	95.4	432	15	US-10-284-962-5	Sequence 5, Appli
15	2188	95.0	436	14	US-10-156-660-2	Sequence 2, Appli
16	2171	94.3	432	15	US-10-601-319-10	Sequence 10, Appli
17	2165	94.0	432	14	US-10-021-723A-16	Sequence 16, Appli
18	2141	93.0	430	9	US-09-866-379-10	Sequence 10, Appli
19	2070	89.9	412	14	US-10-334-672-1	Sequence 1, Appli
20	2070	89.9	412	14	US-10-334-671-1	Sequence 1, Appli
21	2066	89.7	410	14	US-10-021-723A-13	Sequence 13, Appli
22	2066	89.7	410	14	US-10-021-723A-15	Sequence 15, Appli
23	960.5	41.7	441	14	US-10-021-723A-12	Sequence 12, Appli
24	943.5	41.0	441	15	US-10-282-122A-77792	Sequence 77792, A
25	935.5	40.6	441	14	US-10-021-723A-2	Sequence 2, Appli
26	932.5	40.5	420	14	US-10-021-723A-4	Sequence 4, Appli
27	760	33.0	476	14	US-10-021-723A-10	Sequence 10, Appli
28	611.5	26.6	318	14	US-10-021-723A-14	Sequence 14, Appli
29	476.5	20.7	409	14	US-10-021-723A-8	Sequence 8, Appli
30	472.5	20.5	421	14	US-10-021-723A-6	Sequence 6, Appli
31	195	8.5	426	15	US-10-257-174-44	Sequence 44, Appli
32	195	8.5	426	15	US-10-343-357-6	Sequence 6, Appli
33	121	5.3	465	16	US-10-662-914-10	Sequence 10, Appli
34	119	5.2	439	14	US-10-062-848-3	Sequence 3, Appli
35	119	5.2	439	16	US-10-776-104-3	Sequence 10, Appli
36	119	5.2	440	15	US-10-442-538-106	Sequence 106, App
37	119	5.2	440	15	US-10-442-538-128	Sequence 128, App
38	119	5.2	440	15	US-10-442-538-151	Sequence 151, App
39	119	5.2	449	14	US-10-062-848-12	Sequence 12, Appli
40	119	5.2	449	16	US-10-776-104-12	Sequence 12, Appli
41	119	5.2	465	13	US-10-083-452-8	Sequence 8, Appli
42	119	5.2	465	14	US-10-062-848-78	Sequence 78, Appli
43	119	5.2	465	14	US-10-229-358-6	Sequence 6, Appli
44	119	5.2	465	16	US-10-662-914-4	Sequence 4, Appli
45	119	5.2	465	16	US-10-776-104-78	Sequence 78, Appli

ALIGNMENTS

RESULT 1
US-09-777-566A-2
; Sequence 2, Application US/09777566A
; Patent No. US20010055788A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-6
; CURRENT APPLICATION NUMBER: US/09/777,566A
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n is any nucleotide
US-09-777-566A-2

Query Match 100.0%; Score 2302; DB 9; Length 440;
Best Local Similarity 100.0%; Pred. NO. 28-193;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 60
Db 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 60
Qy 61 TWPVKLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVVAIIADVDERTRKTGE 120
Db 61 TWPVKLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVVAIIADVDERTRKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
Qy 181 ROTAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 ROTAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLQQAQMPGPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSATPLLDLIMAA 300
Db 241 EIFLLQQAQMPGPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSATPLLDLIMAA 300
Qy 301 LTPHPPQKQAYGVTLPSTVLFIAGHDTNLNLGGALELNLWTLPGQPDNTPPGGELVFERW 360
Db 301 LTPHPPQKQAYGVTLPSTVLFIAGHDTNLNLGGALELNLWTLPGQPDNTPPGGELVFERW 360
Qy 361 RRLSDNSQWIQVSLVFTQLOMRDKTPLSLNTPPGEVKLTLAGEERNAQMGCSLAGFTQ 420
Db 361 RRLSDNSQWIQVSLVFTQLOMRDKTPLSLNTPPGEVKLTLAGEERNAQMGCSLAGFTQ 420
Qy 421 IVNEARIPACSLRSHHHHH 440
Db 421 IVNEARIPACSLRSHHHHH 440

RESULT 2
US-09-866-379-2
; Sequence 2, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION.
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-866-379-2

Query Match 100.0%; Score 2302; DB 9; Length 440;
Best Local Similarity 100.0%; Pred. No. 2e-193;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 60

Db 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 60
Qy 61 TWPVKLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVVAIIADVDERTRKTGE 120
Db 61 TWPVKLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVVAIIADVDERTRKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
Qy 181 ROTAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 ROTAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLQQAQMPGPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSATPLLDLIMAA 300
Db 241 EIFLLQQAQMPGPGWGRITDSHOWNTLLSLHNAQFYLLQRTPEVARSATPLLDLIMAA 300
Qy 301 LTPHPPQKQAYGVTLPSTVLFIAGHDTNLNLGGALELNLWTLPGQPDNTPPGGELVFERW 360
Db 301 LTPHPPQKQAYGVTLPSTVLFIAGHDTNLNLGGALELNLWTLPGQPDNTPPGGELVFERW 360
Qy 361 RRLSDNSQWIQVSLVFTQLOMRDKTPLSLNTPPGEVKLTLAGEERNAQMGCSLAGFTQ 420
Db 361 RRLSDNSQWIQVSLVFTQLOMRDKTPLSLNTPPGEVKLTLAGEERNAQMGCSLAGFTQ 420
Qy 421 IVNEARIPACSLRSHHHHH 440
Db 421 IVNEARIPACSLRSHHHHH 440

RESULT 3
US-10-034-985-2
; Sequence 2, Application US/10034985
; Publication No. US20030049815A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/034,985
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-034-985-2

Query Match 100.0%; Score 2302; DB 14; Length 440;
Best Local Similarity 100.0%; Pred. No. 2e-193;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 60
Db 1 MKAILPFLSLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 60
Qy 61 TWPVKLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVVAIIADVDERTRKTGE 120
Db 61 TWPVKLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVVAIIADVDERTRKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180

```
QY 181 ROTAFRELERVLPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERVLPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOAQOQMPFGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLDLIMAA 300
DB 241 EIFLLOAQOQMPFGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLDLIMAA 300
QY 301 LTPHPKQKAYGVTLPSTVLFIAGHDNTNLNLCALPSELKVSADNVSLTGAVSLASMLT 360
DB 301 LTPHPKQKAYGVTLPSTVLFIAGHDNTNLNLCALPSELKVSADNVSLTGAVSLASMLT 360
QY 361 RRLSDNSQWTOVSLVFTQLOQMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
DB 361 RRLSDNSQWTOVSLVFTQLOQMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
QY 421 IVNEARIPACSLRSHHHHH 440
DB 421 IVNEARIPACSLRSHHHHH 440

RESULT 4
US-10-430-356-2
; Sequence 2, Application US/10430356
; Publication No. US20030232041A1
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/430,356
; CURRENT FILING DATE: 2003-05-05
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-430-356-2
```

```
Query Match 100.0%; Score 2302; DB 15; Length 440;
Best Local Similarity 100.0%; Pred. No. 2e-193;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAILIPFLLSLLIPTQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMQDVTDPWP 60
DB 1 MKAILIPFLLSLLIPTQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMQDVTDPWP 60
QY 61 TWPVKGLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
DB 61 TWPVKGLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
QY 181 ROTAFRELERVLPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERVLPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOAQOQMPFGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLDLIMAA 300
DB 241 EIFLLOAQOQMPFGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLDLIMAA 300
```

```
QY 301 LTPHPKQKAYGVTLPSTVLFIAGHDNTNLNLCALPSELKVSADNVSLTGAVSLASMLT 360
DB 301 LTPHPKQKAYGVTLPSTVLFIAGHDNTNLNLCALPSELKVSADNVSLTGAVSLASMLT 360
QY 361 RRLSDNSQWTOVSLVFTQLOQMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
DB 361 RRLSDNSQWTOVSLVFTQLOQMRDKTPLSLNTPPGEVKLTLAGCEERNAQCMCSLAGFTQ 420
QY 421 IVNEARIPACSLRSHHHHH 440
DB 421 IVNEARIPACSLRSHHHHH 440

RESULT 5
US-10-601-319-2
; Sequence 2, Application US/10601319
; Publication No. US20040091968A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith A.
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson Robert
; APPLICANT: Garrett, James B.
; APPLICANT: O' Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: RECOMBINANT PHYTASES AND METHODS OF MAKING
; FILE REFERENCE: 09010-029011
; CURRENT APPLICATION NUMBER: US/10/601,319
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Escherichia coli
US-10-601-319-2
```

```
Query Match 100.0%; Score 2302; DB 15; Length 440;
Best Local Similarity 100.0%; Pred. No. 2e-193;
Matches 440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKAILIPFLLSLLIPTQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMQDVTDPWP 60
DB 1 MKAILIPFLLSLLIPTQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMQDVTDPWP 60
QY 61 TWPVKGLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
DB 61 TWPVKGLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
QY 181 ROTAFRELERVLPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELERVLPQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOAQOQMPFGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLDLIMAA 300
DB 241 EIFLLOAQOQMPFGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRATPLDLIMAA 300
```



```
QY 301 LTPHPKQAYGVTLPSTVLFIAHDTNLANLGALNLTLPQPDNTTPGGELVFERW 360
DB 301 LTPHPKQAYGVTLPSTVLFIAHDTNLANLGALNLTLPQPDNTTPGGELVFERW 360
QY 361 RRLSDNSQWIQSVLFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
DB 361 RRLSDNSQWIQSVLFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
QY 421 IVNEARIPACSLRSHHHH 440
DB 421 IVNEARIPACSLRSHHHH 440

RESULT 6
US-09-866-379-8
; Sequence 8, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-866-379-8

Query Match 97.1%; Score 2235; DB 9; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.5e-187;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPWP 60
DB 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPWP 60
QY 61 TWPVKLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPCQGVAIADVDERTRKTGE 120
DB 61 TWPVKLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPCQGVAIADVDERTRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
QY 181 ROTAFRELRLVNPQSNCLKREKODESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELRLVNPQSNCLKREKODESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLQOACQMPGPGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
DB 241 EIFLQOACQMPGPGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
QY 301 LTPHPKQAYGVTLPSTVLFIAHDTNLANLGALNLTLPQPDNTTPGGELVFERW 360
```

```
DB 301 LTPHPKQAYGVTLPSTVLFIAHDTNLANLGALNLTLPQPDNTTPGGELVFERW 360
QY 361 RRLSDNSQWIQSVLFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
DB 361 RRLSDNSQWIQSVLFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 7
US-10-156-660-4
; Sequence 4, Application US/10156660
; Publication No. US20030103958A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson R.
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; APPLICANT: Mathur, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM
; FILE REFERENCE: 09010-029007
; CURRENT APPLICATION NUMBER: US/10/156,660
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-156-660-4

Query Match 97.1%; Score 2235; DB 14; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.5e-187;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPWP 60
DB 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPWP 60
QY 61 TWPVKLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPCQGVAIADVDERTRKTGE 120
DB 61 TWPVKLWLTTPRGELIAYLGHYQORLVADGLLAKKGCPCQGVAIADVDERTRKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
QY 181 ROTAFRELRLVNPQSNCLKREKODESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELRLVNPQSNCLKREKODESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLQOACQMPGPGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
DB 241 EIFLQOACQMPGPGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
```

Qy	301	LTPHPQKQAYGVTLPTSVLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGGELVFERW	360
Db	301	LTPHPQKQAYGVTLPTSVLFIAGHDTNLNLGGALELNWTLPGQPDNTPPGGELVFERW	360
Qy	361	RLSDNSQWIOVSLVFQTLQWRDKTPLSLNTPPGCEVKLTLAGCERNAQGCWCSLAGFTQ	420
Db	361	RLSDNSQWIOVSLVFQTLQWRDKTPLSLNTPPGCEVKLTLAGCERNAQGCWCSLAGFTQ	420
Qy	421	IVNEARIPACSL 432	
Db	421	IVNEARIPACSL 432	
RESULT 8			
US-10-282-122A-43351			
; Sequence 43351, Application US/10282122A			
; Publication No. US20040029129A1			
; GENERAL INFORMATION:			
; APPLICANT: Wang, Liangsu			
; APPLICANT: Zamudio, Carlos			
; APPLICANT: Malone, Cheryl			
; APPLICANT: Haselbeck, Robert			
; APPLICANT: Ohlsen, Kari			
; APPLICANT: Zyekind, Judith			
; APPLICANT: Wall, Daniel			
; APPLICANT: Trawick, John			
; APPLICANT: Carr, Grant			
; APPLICANT: Yamamoto, Robert			
; APPLICANT: Forsyth, R.			
; APPLICANT: Xu, H.			
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms			
; FILE REFERENCE: ELITRA.034A			
; CURRENT APPLICATION NUMBER: US/10/282,122A			
; CURRENT FILING DATE: 2003-02-20			
; PRIOR APPLICATION NUMBER: 60/191,078			
; PRIOR FILING DATE: 2000-03-21			
; PRIOR APPLICATION NUMBER: 60/206,848			
; PRIOR FILING DATE: 2000-05-23			
; PRIOR APPLICATION NUMBER: 60/207,727			
; PRIOR FILING DATE: 2000-05-26			
; PRIOR APPLICATION NUMBER: 60/230,335			
; PRIOR FILING DATE: 2000-09-06			
; PRIOR APPLICATION NUMBER: 60/230,347			
; PRIOR FILING DATE: 2000-09-09			
; PRIOR APPLICATION NUMBER: 60/242,578			
; PRIOR FILING DATE: 2000-10-23			
; PRIOR APPLICATION NUMBER: 60/253,625			
; PRIOR FILING DATE: 2000-11-27			
; PRIOR APPLICATION NUMBER: 60/257,931			
; PRIOR FILING DATE: 2000-12-22			
; PRIOR APPLICATION NUMBER: 60/267,636			
; PRIOR FILING DATE: 2001-02-09			
; PRIOR APPLICATION NUMBER: 60/269,308			
; PRIOR FILING DATE: 2001-02-16			
; Remaining Prior Application data removed - See File Wrapper or PALM.			
; NUMBER OF SEQ ID NOS: 78614			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 43351			
; LENGTH: 432			
; TYPE: PRT			
; ORGANISM: Escherichia coli			
US-10-282-122A-43351			
Query Match 97.1%; Score 2235; DB 15; Length 432;			
Best Local Similarity 99.5%; Pred. No. 1.5e-187;			
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0			
Qy	1	MKAILIFLSLLIPLTPQSAFQSEPELKLSVIVSRHGVRAPTKATQLMODVTPDAMP	60
Db	1	MKAILIFLSLLIPLTPQSAFQSEPELKLSVIVSRHGVRAPTKATQLMODVTPDAMP	60
Qy	61	TWPVKLWLTTPRGELIAYLGHYQQRIVADGLLAKKCCPSGQVAIIADVDRTRKTGE	120

```
Db 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQDLDNANVTDAILSRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQDLDNANVTDAILSRAGGSIAFTGH 180
Qy 181 RQAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLOAQOGMPGPGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFLLOAQOGMPGPGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Qy 301 LTPHPPOKQAYGVTLPSTVFLFIAGHDTNLANLGALLENWTLPGQPDNTPPGGELVFERW 360
Db 301 LTPHPPOKQAYGVTLPSTVFLFIAGHDTNLANLGALLENWTLPGQPDNTPPGGELVFERW 360
Qy 361 RRLSDNSQWIVQSVLFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQWIVQSVLFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 10
US-10-472-317-41
; Sequence 41, Application US/10472317
; Publication No. US20040185562A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Incorporated
; TITLE OF INVENTION: Myo-Inositol Oxygenases
; FILE REFERENCE: 10829/003US1
; CURRENT APPLICATION NUMBER: US/10/472,317
; CURRENT FILING DATE: 2003-09-17
; PRIOR APPLICATION NUMBER: PCT/US02/08404
; PRIOR FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: US 60/277,148
; PRIOR FILING DATE: 2001-03-19
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-472-317-41
```

```
Query Match 97.1%; Score 2235; DB 16; Length 432;
Best Local Similarity 99.5%; Pred. No. 1.5e-187;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLIPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTPDWP 60
Db 1 MKAILPFLSLIPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTPDWP 60
Qy 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
Db 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQDLDNANVTDAILSRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQDLDNANVTDAILSRAGGSIAFTGH 180
Qy 181 RQAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLOAQOGMPGPGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFLLOAQOGMPGPGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
```

```
Qy 301 LTPHPPOKQAYGVTLPSTVFLFIAGHDTNLANLGALLENWTLPGQPDNTPPGGELVFERW 360
Db 301 LTPHPPOKQAYGVTLPSTVFLFIAGHDTNLANLGALLENWTLPGQPDNTPPGGELVFERW 360
Qy 361 RRLSDNSQWIVQSVLFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQWIVQSVLFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 11
US-10-284-962-3
; Sequence 3, Application US/10284962
; Publication No. US20030206913A1
; GENERAL INFORMATION:
; APPLICANT: Weibel, Donald E.
; APPLICANT: Orr, Donald E.
; APPLICANT: Ruch, Frank E.
; TITLE OF INVENTION: PHYTASE-CONTAINING ANIMAL FOOD AND METHOD
; FILE REFERENCE: 834460-71725
; CURRENT APPLICATION NUMBER: US/10/284,962
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/335,303
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 432
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-284-962-3
```

```
Query Match 96.5%; Score 2221; DB 15; Length 432;
Best Local Similarity 99.1%; Pred. No. 2.6e-186;
Matches 428; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLIPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTPDWP 60
Db 1 MKAILPFLSLIPLTPQSAFAQSEPELKESVIVSRHGVRAPTKATQLMQDVTPDWP 60
Qy 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
Db 61 TWPVKLGWLTTPRGGLIAYLGHYQRLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQDLDNANVTDAILSRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCQDLDNANVTDAILSRAGGSIAFTGH 180
Qy 181 RQAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Db 181 RQAFRELERVLNFPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
Qy 241 EIFLLOAQOGMPGPGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFLLOAQOGMPGPGWGRITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Qy 301 LTPHPPOKQAYGVTLPSTVFLFIAGHDTNLANLGALLENWTLPGQPDNTPPGGELVFERW 360
Db 301 LTPHPPOKQAYGVTLPSTVFLFIAGHDTNLANLGALLENWTLPGQPDNTPPGGELVFERW 360
Qy 361 RRLSDNSQWIVQSVLFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Db 361 RRLSDNSQWIVQSVLFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432
```

RESULT 12

US-10-284-962-14
; Sequence 14, Application US/10284962
; Publication No. US20030206913A1
; GENERAL INFORMATION:
; APPLICANT: Weibel, Douglas M.
; APPLICANT: Orr, Donald E.
; APPLICANT: Ruch, Frank E.
; TITLE OF INVENTION: PHYTASE-CONTAINING ANIMAL FOOD AND METHOD
; FILE REFERENCE: 834460-71725
; CURRENT APPLICATION NUMBER: US/10/284,962
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/335,303
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 432
; TYPE: PRN
; ORGANISM: Escherichia coli
US-10-284-962-14

Query Match 96.5%; Score 2221; DB 15; Length 432;
Best Local Similarity 99.1%; Pred. No. 2.6e-186;
Matches 428; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKAILIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
Db 1 MKAILIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60

Qy 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVVAIADVDERTRKTGE 120
Db 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVVAIADVDERTRKTGE 120

Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGQCLDNANVTDAIISRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGQCLDNANVTDAIISRAGGSIAFTGH 180

Qy 181 RQTAFRELERVLNFPQSNLCIKREKQDESLTQALPSELKVSADNVSILTGAVSLASMLT 240
Db 181 RQTAFRELERVLNFPQSNLCIKREKQDESLTQALPSELKVSADNVSILTGAVSLASMLT 240

Qy 241 EIFLLQQAQMPGPGWGRITDSHWNLTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFLLQQAQMPGPGWGRITDSHWNLTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALSLNWTLPQPDNTPPGGELVFERW 360
Db 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALSLNWTLPQPDNTPPGGELVFERW 360

Qy 361 RRLSDNSQWIVQSVLFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420
Db 361 RRLSDNSQWIVQSVLFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420

Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 13

US-10-266-041-1
; Sequence 1, Application US/10266041
; Publication No. US20030072844A1
; GENERAL INFORMATION:
; APPLICANT: Lei, Xingen
; TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
; FILE REFERENCE: 19603/2791
; CURRENT APPLICATION NUMBER: US/10/266,041
; CURRENT FILING DATE: 2002-10-07
; PRIOR APPLICATION NUMBER: US/09/540,149
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 60/127,032
; PRIOR FILING DATE: 1999-03-31

; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 433
; TYPE: PRN
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (433)
; OTHER INFORMATION: Xaa at position 433 in this sequence is unknown
US-10-266-041-1

Query Match 96.5%; Score 2221; DB 14; Length 433;
Best Local Similarity 99.1%; Pred. No. 2.6e-186;
Matches 428; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKAILIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
Db 1 MKAILIPFLSLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60

Qy 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVVAIADVDERTRKTGE 120
Db 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVVAIADVDERTRKTGE 120

Qy 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGQCLDNANVTDAIISRAGGSIAFTGH 180
Db 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGQCLDNANVTDAIISRAGGSIAFTGH 180

Qy 181 RQTAFRELERVLNFPQSNLCIKREKQDESLTQALPSELKVSADNVSILTGAVSLASMLT 240
Db 181 RQTAFRELERVLNFPQSNLCIKREKQDESLTQALPSELKVSADNVSILTGAVSLASMLT 240

Qy 241 EIFLLQQAQMPGPGWGRITDSHWNLTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
Db 241 EIFLLQQAQMPGPGWGRITDSHWNLTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300

Qy 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALSLNWTLPQPDNTPPGGELVFERW 360
Db 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALSLNWTLPQPDNTPPGGELVFERW 360

Qy 361 RRLSDNSQWIVQSVLFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420
Db 361 RRLSDNSQWIVQSVLFTQLOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420

Qy 421 IVNEARIPACSL 432
Db 421 IVNEARIPACSL 432

RESULT 14

US-10-284-962-5
; Sequence 5, Application US/10284962
; Publication No. US20030206913A1
; GENERAL INFORMATION:
; APPLICANT: Weibel, Douglas M.
; APPLICANT: Orr, Donald E.
; APPLICANT: Ruch, Frank E.
; TITLE OF INVENTION: PHYTASE-CONTAINING ANIMAL FOOD AND METHOD
; FILE REFERENCE: 834460-71725
; CURRENT APPLICATION NUMBER: US/10/284,962
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: US 60/335,303
; PRIOR FILING DATE: 2001-10-31
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 432
; TYPE: PRN
; ORGANISM: Escherichia coli
US-10-284-962-5

Query Match 95.4%; Score 2197; DB 15; Length 432;
Best Local Similarity 98.1%; Pred. No. 3.3e-184;

Matches	424;	Conservative	2;	Mismatches	6;	Indels	0;	Gaps	0;
Qy	1	MKAILIFL	SLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDAMP	60					
Db	1	MKAILIFL	SLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQMLQMDVTPDAMP	60					
Qy	61	TWPVKL	GLWLTTPRGGELIAYLGHYQORLIVADGLLAKGCPQGOVAIIADVDETRKTGE	120					
Db	61	TWPVKL	GLWLTTPRGGELIAYLGHYQORLIVADGLLAKGCPQGOVAIIADVDETRKTGE	120					
Qy	121	AFAAGL	APDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIADFTGH	180					
Db	121	AFAAGL	APDCAITVHTQADTSSPDPLFPLKTVGCOLDNANVTDAILSRAGGSIADFTGH	180					
Qy	181	QTAFRE	LVLNFPQSNILCKRKQDESCSLTOALPSELKVSADNYSLTCAVSLASMLT	240					
Db	181	QTAFRE	LVLNFPQSNILNLRKQESCNLTQALPSELKVSADNYSLTCAVSLASMLT	240					
Qy	241	EIFLLQ	AQGMPEPGWGRITDSHQWNTILSLHNAQFYLLQRTPEVARSRATPLLDLIMAA	300					
Db	241	EIFLLQ	AQGMPEPGWGRITDSHQWNTILSLHNAQFYLLQRTPEVARSRATPLLDLIKTA	300					
Qy	301	LTHPPQ	KQAYGVTLPTSVLFIAGHDNTNLNLGGALEHLNWTLPQPDNTPPGGLVLFERW	360					
Db	301	LTHPPQ	KQAYGVTLPTSVLFIAGHDNLNLNLGGALEHLNWTLPQPDNTPPGGLVLFERW	360					
Qy	361	RLSDNS	QWLTQVSLVFOTLOQMURDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ	420					
Db	361	RLSDNS	QWLTQVSLVFOTLQOMURDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ	420					
Qy	421	IVNEAR	IPACSL 432						
Db	421	IVNEAR	IPACSL 432						

RESULT 15

```

US-10-156-660-2
; Sequence 2, Application US/10156660
; Publication No. US20030103958A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; APPLICANT: Kretz, Keith
; APPLICANT: Gray, Kevin A.
; APPLICANT: Barton, Nelson R.
; APPLICANT: Garrett, James B.
; APPLICANT: O'Donoghue, Eileen
; APPLICANT: Machur, Eric J.
; TITLE OF INVENTION: PHYTASES, NUCLEIC ACIDS ENCODING THEM
; FILE REFERENCE: 09010-029007
; CURRENT APPLICATION NUMBER: US/10/156,660
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: US 09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 436
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: modified phytase enzyme
US-10-156-660-2

```

Query Match	95.0%;	Score 2188;	DB 14;	Length 436;
Best Local Similarity	97.7%;	Pred. No. 2.1e-183;		
Matches 425;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0;
Qy	1	MKAILIPLFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTTPDAMP	60	
Db	1	MKAILIPLFLSLLIPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTTPDAMP	60	
Qy	61	TWPVKLGLWLPTRGGELIAYIGHYORQLVADGLLAKKGCPOSGQVAIIADYDERTKTCGE	120	
Db	61	TWPVKLGELTPRGGELIAYIGHYWRQLVADGLLPKGCPOSGQVAIIADYDERTKTCGE	120	
Qy	121	AFAAGLAPDCAIIVHTQADTSSPDPLFNPLKTVGCQLDANVNTDAILSRAGSIADETGH	180	
Db	121	AFAAGLAPDCAIIVHTQADTSSPDPLFNPLKTVGCQLDANVNTDALLERAGSIADETGH	180	
Qy	181	QYTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT	240	
Db	181	YQTAFRELERVLNPPQSNLCLKREKQDESCSLTQALPSELKVSADCVSLTGAVSLASMLT	240	
Qy	241	EIFLLQQAQGMPEGCWGRITDISHQWNTLLSHNAQFVLLQRTPEVARSRAPFLDLIMAA	300	
Db	241	EIFLLQQAQGMPEGCWGRITDISHQWNTLLSHNAQFVLLQRTPEVARSRAPFLDLIKTA	300	
Qy	301	LTPHPPQKQAVGTVLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVLFRRW	360	
Db	301	LTPHPPQKQAVGTVLPTSVLFIAGHDTNLANLGGALELNWTLPGQPDNTPPGGELVLFRRW	360	
Qy	361	RLSDNSQWQVSVLFTLOQMQRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ	420	
Db	361	RLSDNSQWQVSVLFTLOQMQRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ	420	
Qy	421	IVNEARIPACSLRSH	435	
Db	421	IVNEARIPACSLRSH	435	

Search completed: March 11, 2005, 23:20:01
Job time : 144 secs.

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 22:58:45 ; Search time 41 Seconds
(without alignments)
1032.570 Million cell updates/sec

Title: US-10-034-985-2

Perfect score: 2302

Sequence: 1 MKAILPFLSLIPLTPQSA.....IVNEARIPACSLRSHHHHH 440

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2235	97.1	432	2 B36733	acid phosphatase (
2	2210	96.0	434	2 H90770	phosphoanhydride p
3	2210	96.0	444	2 D85633	hypothetical prote
4	943.5	41.0	441	2 AC0201	acid phosphatase (
5	545	23.7	413	2 F90773	periplasmic glucos
6	545	23.7	413	2 B85636	periplasmic glucos
7	541	23.5	417	2 S25627	glucose-1-phosphat
8	539	23.4	413	2 JV0087	glucose-1-phosphat
9	526	22.8	414	2 E87316	periplasmic phosph
10	524	22.8	413	2 AG0632	glucose-1-phosphat
11	139.5	6.1	423	1 S06167	acid phosphatase (
12	133	5.8	416	2 T16058	hypothetical prote
13	131	5.7	423	2 A33395	acid phosphatase (
14	122.5	5.3	421	2 S14742	acid phosphatase (
15	117	5.1	386	1 JH0610	acid phosphatase (
16	113.5	4.9	344	2 B89130	protein F52E1.8 [i
17	109	4.7	2554	1 TVFF7L	kinase-related pro
18	108.5	4.7	10223	2 T30225	polyketide synthas
19	107.5	4.7	457	2 A56925	paired box transcr
20	107	4.6	450	2 A54429	paired box transcr
21	106.5	4.6	1048	1 BVESGC	exonuclease (EC 3.
22	105.5	4.6	459	2 S52250	paired box transcr
23	104.5	4.5	2535	2 AC0304	probable hemolysin
24	102	4.4	638	2 A29440	signal recognition
25	101.5	4.4	397	2 C81716	hypothetical prote
26	101.5	4.4	479	1 JN0890	acid phosphatase (
27	101.5	4.4	524	2 H75530	probable acid-CoA
28	101.5	4.4	1047	2 C85535	ATP-dependent dsDN
29	101.5	4.4	1047	2 G90684	ATP-dependent dsDN

ALIGNMENTS

RESULT 1

B36733

acid phosphatase (EC 3.1.3.2) precursor - Escherichia coli (strain K-12)

N;Alternate names: phosphoanhydride phosphohydrolase, periplasmic; phytase 6

C;Species: Escherichia coli

C;Date: 19-Apr-1991 #sequence revision 19-Apr-1991 #text change 09-Jul-2004

C;Accession: B36733; S18018; B64839; A26534; S17960; S33278

R;Dassa, J.; Marck, C.; Boquet, P.L.

J. Bacteriol. 172, 5497-5500, 1990

A;Title: The complete nucleotide sequence of the Escherichia coli gene appA reveals sign

A;Reference number: A36733; MUID:90368616; PMID:2168385

A;Accession: B36733

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-432 <DAS>

A;Cross-references: UNIPROT:P07102; GB:M58708; NID:g145283; PIDN:AAA72086.1; PID:g145283

R;Greiner, R.; Jany, K.D.

Biol. Chem. Hoppe-Seyler 372, 664-665, 1991

A;Title: Characterization of a phytase from Escherichia coli.

A;Reference number: S18018

A;Accession: S18018

A;Molecule type: protein

A;Residues: 23-33 <GRE>

R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A;Title: The complete genome sequence of Escherichia coli K-12.

A;Reference number: A64720; MUID:97426617; PMID:9278503

A;Accession: B64839

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-432 <BLAT>

A;Cross-references: GB:AE000200; GB:U00096; NID:g2367111; PIDN:AACT4065.1; PID:g1787215;

A;Experimental source: strain K-12, substrain MG1655

R;Touati, E.; Danchin, A.

Biochimie 69, 215-221, 1987

A;Title: The structure of the promoter and amino terminal region of the pH 2.5 acid phosph

A;Reference number: A26534; MUID:87271766; PMID:3038201

A;Accession: A26534

A;Molecule type: DNA

A;Residues: 1-50, 'NAGCHPRRMANLAG', 65, 'T', 67-74, 'DV', 77-111, 'S', <TOU>

A;Cross-references: GB:X05471; NID:g40925; PIDN:CAA29031.1; PID:g40927

R;Dassa, J.; Fahb, H.; Marck, C.; Dion, M.; Kleffer-Bontemps, M.; Boquet, P.L.

Mol. Gen. Genet. 229, 341-352, 1991

A;Title: A new oxygen-regulated operon in Escherichia coli comprises the genes for a put

A;Reference number: S17958; MUID:92049231; PMID:1658595

A;Accession: S17960

A;Status: not compared with conceptual translation

A;Molecule type: DNA

A;Residues: 1-17 <DA2>

A;Cross-references: GB:S63811; NID:g238656; PIDN:AAB20286.1; PID:g238659

R;Greiner, R.; Konietzny, U.; Jany, K.D.

Arch. Biochem. Biophys. 303, 107-113, 1993
A;Title: Purification and characterization of two phytases from *Escherichia coli*.
A;Reference number: S33278; MUID:9325656; PMID:8387749
A;Accession: S33278
A;Molecule type: protein
A;Residues: 23-31, A, 33-34 <GR2>
C;Comment: In addition to CAMP-mediated control, this enzyme is induced when bacterial c
naerobic conditions.
C;Genetics:
A;Gene: appA
A;Map position: 25 min
C;Keywords: monomer; periplasmic space; phosphohistidine; phosphoprotein; phosphoric mon
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-112/Product: acid phosphatase #status predicted <MAT>
F;39/Active site: Arg #status predicted
F;39/Active site: His (phosphohistidine intermediate) #status predicted

Query Match 97.1%; Score 2235; DB 2; Length 432;
Best Local Similarity 99.5%; Pred. No. 6.1e-171;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY . 1 MKAIIIPFLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPWP 60
DB 1 MKAIIIPFLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPWP 60
QY 61 TWPVKLGWLTGRGGLIAYLGHYQRLVADGLLAKKGCQSQGVAIADVDERTKTGE 120
DB 61 TWPVKLGWLTGRGGLIAYLGHYQRLVADGLLAKKGCQSQGVAIADVDERTKTGE 120
QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
QY 181 ROTAFRELRLVLPFQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 ROTAFRELRLVLPFQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
QY 241 EIFLLOQAQGMPEPGWGRITDTSQWNTLLSHNAQFVLLQRTPEVARSRATPLLDLIMAA 300
DB 241 EIFLLOQAQGMPEPGWGRITDTSQWNTLLSHNAQFVLLQRTPEVARSRATPLLDLIMAA 300
QY 301 LTPHPKQKQAYGVTLPTSVLFIAGHDNLANLGALNELNWTLPQGPDPNTPPGGELVPERW 360
DB 301 LTPHPKQKQAYGVTLPTSVLFIAGHDNLANLGALNELNWTLPQGPDPNTPPGGELVPERW 360
QY 361 RLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
DB 361 RLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
QY 421 IVNEARIPACSL 432
DB 421 IVNEARIPACSL 432

RESULT 2
H90770
phosphoanhydride phosphorylase [imported] - *Escherichia coli* (strain O157:H7, substrain
C;Species: *Escherichia coli*
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: H90770
R;Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: H90770
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-434 <HAY>
A;Cross-references: UNIPROT:Q8XC29; GB:BA000007; PIDN:BA034559.1; PID:g13360596; GSPDB:G
A;Experimental source: strain O157:H7, substrain RMD 050952
C;Genetics:
A;Gene: ECa1136

Query Match 96.0%; Score 2210; DB 2; Length 434;
Best Local Similarity 98.4%; Pred. No. 6.1e-169;
Matches 427; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
QY 1 MKAIIIPFLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPDA 58
DB 1 MKAIIIPFLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPDA 60
QY 59 WPTWPKLGWLTGRGGLIAYLGHYQRLVADGLLAKKGCQSQGVAIADVDERTKT 118
DB 61 WPNWPKLGWLTGRGGLIAYLGHYQRLVADGLLTKKGCQSQGVAIADVDERTKT 120
QY 119 GEAFAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFT 178
DB 121 GEAFAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFT 180
QY 179 GHRQTAFRELRLVLPFQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASML 238
DB 181 GHRQTAFRELRLVLPFQSNCLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASML 240
QY 239 LTFEILLOQAQGMPEPGWGRITDTSQWNTLLSHNAQFVLLQRTPEVARSRATPLLDLIM 298
DB 241 LTFEILLOQAQGMPEPGWGRITDTSQWNTLLSHNAQFVLLQRTPEVARSRATPLLDLIM 300
QY 299 AALTTPHPKQKQAYGVTLPTSVLFIAGHDNLANLGALNELNWTLPQGPDPNTPPGGELVPE 358
DB 301 IALTTPHPKQKQAYGVTLPTSVLFIAGHDNLANLGALNELNWTLPQGPDPNTPPGGELVPE 360
QY 359 RWRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGF 418
DB 361 RWRLSDNSQWIOVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGF 420
QY 419 TQIVNEARIPACSL 432
DB 421 TQIVNEARIPACSL 434

RESULT 3
D85633
hypothetical protein appA [imported] - *Escherichia coli* (strain O157:H7, substrain EDL93:
C;Species: *Escherichia coli*
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: D85633
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew,
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimailanta, E.; Potamouisis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: D85633
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-444 <STO>
A;Cross-references: UNIPROT:Q8XC29; GB:AE005174; NID:g12514245; PIDN:AAG55528.1; GSPDB:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: appA

Query Match 96.0%; Score 2210; DB 2; Length 444;
Best Local Similarity 98.4%; Pred. No. 6.3e-169;
Matches 427; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
QY 1 MKAIIIPFLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPDA 58
DB 11 MKAIIIPFLLPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTDPDA 70
QY 59 WPTWPKLGWLTGRGGLIAYLGHYQRLVADGLLAKKGCQSQGVAIADVDERTKT 118
DB 71 WPNWPKLGWLTGRGGLIAYLGHYQRLVADGLLTKKGCQSQGVAIADVDERTKT 130
QY 119 GEAFAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFT 178
DB 131 GEAFAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFT 190

Qy 179 CHROTAFRELERVLPQSNICLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASM 238
Db 191 CHROTAFRELERVLPQSNICLNREKQDESCSLTQALPSELKVSADNVSLTGAVSLASM 250
Qy 239 LTELFLQOQGMPEPGWGRITDSHWNLTLLSHNAQFYLLQRTPEVARSRAPELDDLM 298
Db 251 LTELFLQOQGMPEPGWGRITDSHWNLTLLSHNAQFYLLQRTPEVARSRAPELDDLM 310
Qy 299 AALTPHPQKQAYGVTLPTSVLFIAGHDTNLNLAGGALNLTLPQOPDNTPPCGELVFE 358
Db 311 IALTTPHPQKQAYGVTLPTSVLFIAGHDTNLNLAGGALNLTLPQOPDNTPPCGELVFE 370
Qy 359 RWRLSDNSQWQVSLVFQTLQOQMRDKTPLSLNTPPGEVKLTLAGCBERNAQMGCSLAGF 418
Db 371 RWRLSDNSQWQVSLVFQTLQOQMRDKTPLSLNTPPGEVKLTLAGCBERNAQMGCSLAGF 430
Qy 419 TOIVNEARIPACSL 432
Db 431 TOIVNEARIPACSL 444

RESULT 4
AC0201
acid phosphatase (EC 3.1.3.2) [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AC0201
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AC0201
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-441 <KUR>
A;Cross-references: UNIPROT:Q82FP6; GB:AL590842; PIDN:CAC90470.1; PID:g15979685; GSPDB:G
C;Genetics:
A;Gene: YP01648
C;keywords: phosphoric monoester hydrolase

Query Match 41.0%; Score 943.5; DB 2; Length 441;
Best Local Similarity 45.6%; Pred. No. 1.3e-67;
Matches 199; Conservative 66; Mismatches 156; Indels 15; Gaps 7;

Qy 5 LIPPLSLIPLTPQSAFAQSEPE-LKLESVVIVSRHGVRAPTKATQMLQMDVTPDPAWTP 63
Db 13 LVLMGLSLAAITAPVA---AEPGTYTLERVVILSRHGVRSPTKOTQLMNDVTPDKWPQW 69
Qy 64 VKLGMLTPRGELIAYLGHYORQLVADGLLAKKGCPSQSOVAIIADVDERTKGTGEAFA 123
Db 70 VKAGYLTPRGAEVLTMGGFYDGRSLGLLA-AGCPAEGGVAAQADIDQRTLTGQFL 128
Qy 124 AGLAPDCAITVHTQADTSSPDLFNPFLKTVGCOLDNANVTDAILSRAGGSIAFTGHRQT 183
Db 129 DGVAPEGGLVHNOADLKTKTDLPLHPVEAGVCKLDAQTDKAIEQLGGPLDVTVSQYAK 188
Qy 184 AFRELERVLPQSNICLKREKQDESCSLTQALPSELKVSAD--NVSLTGAVSLASMLTE 241
Db 189 PFAQMGVDLVNFAASPYCKSLQOQKTCDFAFHAFANVNVNKEGPKVTLSGPLALSSTLGE 248
Qy 242 IFLLQOQGMPEPGWGRITDSHWNLTLLSHNAQFYLLQRTPEVARSRAPELDDLM 301
Db 249 IFLLQOQGMPEVAVQBLKGAENVVSLLSHNAQFNLMAKTPYIARHKGTPFLQQIDTAL 308
Qy 302 TPHPQKQAYGVTLPTS---VLFIAGHDTNLNLAGGALNLTLPQOPDNTPPCGELVFE 357
Db 309 T---LQIDAQOKPLISAQNRVLEFLGHDTNINIANIGWLGADWQLPQOPDNTPPCGGLVF 365
Qy 358 ERWRRLSDNSQWQVSLVFQTLQOQMRDKTPLSL-NTTPPGEVKLTLAGCBERNAQMGCSLA 416
Db 419 ERWRRLSDNSQWQVSLVFQTLQOQMRDKTPLSL-NTTPPGEVKLTLAGCBERNAQMGCSLA 416

Db 366 ELWQNPDPNHORYVAVKMFYOTMDQLRNAEKLDLKNPAGIISVAVACENNDDKLCELD 425
Qy 417 GFTQIVNEARIPACSL 432
Db 426 TFQKVKVIEPACHI 441

RESULT 5
F90773
periplasmic glucose-1-phosphatase [imported] - Escherichia coli (strain O157:H7, substra
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: F90773
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F90773
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-413 <HAY>
A;Cross-references: UNIPROT:Q8XBZ6; GB:BA000007; PIDN:BA834581.1; PID:g13360618; GSPDB:G
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs1158

Query Match 23.7%; Score 545; DB 2; Length 413;
Best Local Similarity 32.1%; Pred. No. 8.9e-36;
Matches 135; Conservative 76; Mismatches 180; Indels 30; Gaps 10;

Qy 13 IPLTPQSAFAQSEPE-LKLESVVIVSRHGVRAPE-TRATQMLQMDVTPDPAWTPVKLGWLT 70
Db 13 VVLLASNAQAQTVPEGVQLQQLVLMMSRHLRAPLANNNGSVLEQSTPNKWEWDVPGGOLT 72
Qy 71 PRGELIAYLGHYORQLVADGLLAKKGCPSQSOVAIIADVDERTKGTGEAFAAGLAPDC 130
Db 73 TKGGLVLEVYMGHYMRWELAQGMVKSGCEPPDPTVAYANSLQRTVATQAQFFITGAFPGC 132
Qy 131 AITVHTQADTSSPDLFNPFLKTVGCOLDNANVTDAILSRAGGSIAFTGHRQTAFRELER 190
Db 133 DIPVHQEKMGMTDPTFNPVITD-----DSAAFSEQAFAAMEKLSKL--QLTDSYQLLEK 186
Qy 191 VLNFPQSNICLKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLTEIFLLQOQAG 250
Db 187 IVNPKDSPAC----KEKQCCSLVDGKNTFSKAYQDEPGVSGPLKVGNSLVDADFTLQYVEG 242
Qy 251 MP--EPGWRITDSHWNLTLLSHNAQFYLLQRTPEVARSRAPELDDLMIAALTPHPQK 308
Db 243 FPMQDVANGBIKSDQGMKVLKNGYQDSLFTSPEVARNVAKPLVSYIDKAL----- 295
Qy 309 QAYGVTLPTS---VLFIAGHDTNLNLAGGALNLTLPQOPDNTPPCGELVFERWRRLS 364
Db 296 -----VTDRTSAPKITVLVGHDSNIALSLTALDPKPYQLHDQNERPTPGKIVFQRMWDSK 351
Qy 365 DNSQWQVSLVFQTLQOQMRDKTPLSLNTPPGEVKLTLAGCBERNAQMGCSLAGFTQIVNE 424
Db 352 ANRDLMKIEVYQSAEQLRNADALTQAPAQRTVLELSGC-PIDADGFCPMDKFDVNE 410
Qy 425 A 425
Db 411 A 411

RESULT 6
B85636
periplasmic glucose-1-phosphatase [imported] - Escherichia coli (strain O157:H7, substra
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: B85636
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: B85636
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-413 <STO>
A;cross-references: UNIPROT:Q9XBZ6; GB:AE005174; NID:g12514273; PIDN:AGC55550.1; GSPDR:G
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: agp

Query Match	23.7%	Score	545;	DB	2;	Length	413;
Best Local Similarity	32.1%	Pred.	No. 8.9e-36;				
Matches	135;	Conservative	76;	Mismatches	180;	Indels	30;
Gaps	10;						
QY	13	IPLTPOSAFAQSPPE-LKLESVVIVSRHGVRAP-TRKATOLMODVTPDAWPTWPKVLGWLTT	70				
DB	13	VVLASNAQAOTVPEGYQLOQVLMMRHNLRAPLANNQSVLEQSTNKNKPEWDPGQQLT	72				
QY	71	PRGELIAYLGHYQORQLVADGLLAKKGCPQSQQVAIIADVBERTRKTGEAFAGLAPDC	130				
DB	73	TKGGVLEVYMGHYMRWELABQGMVKSGECPDPPTVTVAYANSLQRTVATAQFFITGAFPGC	132				
QY	131	AITVHTQADTSSPDPFLNPLKTCVGLDNNANVTDAILSRAGGSIADFTGHRQTAPARELR	190				
DB	133	DIPVHQEKMGKTWDPTFNPVITD----DSAAFSEQAVAMEKLSKL--QLTDSYQLLEK	186				
QY	191	VLMFPQSNCLCKREKQDESCSLTOALPSELKVSADNVSLTGAVSLASMLTEIFLLQQAQG	250				
DB	187	IVNVKDSAPAC----KEKQCQSLVDGKNTTSAKYQOEPFGVSGPLKGVNSLVDAFTLQYREG	242				
QY	251	MP--EPCWGRIITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLDLLMAALTTPHPQK	308				
DB	243	FPMDQVANGBIKSDQKWVLSKLNGVQDSLSFTSPEVARNVAKPLVSYDIAL--	295				
QY	309	QAYGVTLPTS---VLFIAGHTNLNLGGALELN-WTLPGQPDNTPPGGELVFERWRRLS	364				
DB	296	----VTDRTSAPKITVLVGHDSNIAISLLTALDFKPYQLHDQNETRTPIGGKIVQRWRDSK	351				
QY	365	DNSQWLOVSLAVFOTLQOMRKDTPLSLNTPPGVEVKLTLAGCEERNAQCMCSLACGTQIVNE	424				
DB	352	ANDRLMKIEVYQSAEQLRNADALITQAPAQRTVLELSGC-PIDADGFCPMDKXDFSVLNE	410				
QY	425	A 425					
DB	411	A 411					

```

RESULT 7
S25627
glucose-1-phosphatase (EC 3.1.3.10) precursor - Providencia rettgeri
C:Species: Providencia rettgeri
C:Date: 25-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 09-Jul-2004
C:Accession: S25627
R:Riccio, M.L.; Chiesurin, A.; Lombardi, G.; Satta, G.
submitted to the EMBL Data Library, September 1992
A:Reference number: S25627
A:Accession: S25627
A:Molecule type: DNA
A:Residues: 1-417 <RIC>
A:Cross-references: UNIPROT:Q52309; EMBL:X69201; NID:945771; PIDN:CAA48288.1; PID:945772
A:Experimental source: strain PV7
C:Genetics:
A:Gene: agp
C:Keywords: periplasmic space; phosphohistidine; phosphoprotein; phosphoric monoester hyd
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-417/Product: glucose-1-phosphatase #status predicted <MAT>
F:42/Active site: His (phosphohistidine intermediate) #status predicted
F:314/Active site: His #status predicted

```

Qy	10	SLLIPLTPQAPQASEPELKLESVVIVSRHGVRAPTKATQLMODVTPDAMPWPVKLGWL	69
Db	13	ALFAPIAPTWASTDNQADMWLDQVLVLSRHLRTPVINTGILTEVTDKKWPMDAKSGYL	72
Qy	70	TPRGELIAYLGHYQORQLVADGLLAKKCPQSQG-VAIADVDERTKTKGFAAGLAP	128
Db	73	TTQGGALEVYMGHVFREWIQONKLLADELCPTSNEIDIVLNTSLQRTIATAQFFAAGAF	132
Qy	129	DCALTVHTQADTSSPDLFNPCLKTGVQOLDNANVTDAILSRAGGSIAFTG--HRQTAPR	186
Db	133	GCKVNIHHQPEIGKMDPVFNPIT-----NGSPFEFKALAAAMDDYLGKLSLKAGYE	184
Qy	187	ELERVLPFQPSNCLGKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLTEIFLLQ	246
Db	185	ELDTVLMIKQSK-----KTDKCNLDSQKSNFIEADKEPGVSGPLKIANSAVDAIDLO	240
Qy	247	QAOGMP--EPGCGRITDSHOWNTLLSLHNAQFYLLIQRTPEVARSRATPLLLDILMAALTPH	304
Db	241	YIEGFPADQVAVGLVDTPEKWKKLNTKNAQOETLFTPKIIAKNVAHPILNVIDKGFVSU	300
Qy	305	PPQQAQGVTLPTSVLFIAGHDTNLANLGGALELN-WTLPGOPDNTPPGGELAVFERWRL	363
Db	301	DKGETA-----KFIFLVGHDSNIASLMSAMDPKPYQLOAQOYERTPIGGKLVFORWTDK	353
Qy	364	SDNSQWIQVSLVFOTLOQMRRKTPLSLNTPGCEVKLTLAGEB-ERNAAQMCSLAGFTQIV	422
Db	354	QTKXDPKMKVEVYQATQDLRDNAYLSLETTPPKHVTLELKCDFVDKN--GYCSWEDQKVM	411
Qy	423	NEA 425	
Db	412	AKA 414	

RESULT a

JV0087
glucose-1-phosphatase (EC 3.1.3.10) precursor - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C;Accession: JV0087; H64841
R;Pradel, E.; Marck, C.; Boquet, P.L.
J. Bacteriol. 172, 802-807, 1990
A;Title: Nucleotide sequence and transcriptional analysis of the Escherichia coli agp gene
A;Reference number: JV0087; MUID: 90130318; PMID: 2153660
A;Accession: JV0087
A;Molecule type: DNA
A;Residues: 1-413 <PRA>
A;Cross-references: UNIPROT:P19926; GB:M33807; NID:G145217; PIDN:AAA23426.1; PID:G145218
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Col
. A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID: 97426617; PMID: 9278503
A;Accession: H64841
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-413 <BLAT>
A;Cross-references: GB:A5000202; GB:U00096; NID:G1787233; PIDN:AAC74087.1; PID:G1787237;
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: agp
A;Map position: 23 min
C;Function:
A;Description: essential for growth in a high-phosphate medium containing glucose-1-phosph
A;Note: Optimal at low pH
C;Keywords: homodimer; periplasmic space; phosphohistidine; phosphoprotein; phosphoric m
F;1-22/Domain: signal sequence #status predicted <SIG>
F;23-413/Product: glucose-1-phosphatase #status predicted <MAT>
F;40/Active site: His (phosphohistidine intermediate) #status predicted
F;311/Active site: His #status predicted

Query Match	Score	DB 2;	Length
Best Local Similarity	23.4%;	539;	413;
Pred.	32.1%;		
	No. 2.7e-35;		

Query Match 23.4%; Score 539; DB 2; Length 413;
Best Local Similarity 32.1%; Pred. No. 2.7e-35;

Query Match 23.4%; Score 539; DB 2; Length 413;
Best Local Similarity 32.1%; Pred. No. 2.7e-35;

RESULT 10

Query Match	22.8%	Score 524;	DB 2;	Length 413;
Best Local Similarity	33.0%;	Pred. No. 4.3e-34;		
Matches 135;	Conservative 65;	Mismatches 185;	Indels 24;	Gaps 10
Qy	22	AQSPE-LKLESVIVSRHGVRAP-TRATQLMQDVTDPDAWTPVYKLGWLTPRGELIAY	79	
Db	22	AQTTPEGYQQVLMRSRHLRAPLANNGVLAQSTPNAPWADVPGGQLTKGGVLEVI	81	
Qy	80	LGHYORQLVADGELLAKKGCPOSGOVAIADVDSERTRKTGEAAGLAPDCATAIVHTQAD	139	
Db	82	MGHYTRSWLVAQGLIPSGCEPAPDTVYAYANSLQRTVATAQFFITSAPGCDIPVHHQEK	141	
Qy	140	TSSPDPLNPLKTCWGCLDNANVTDAILSRAGGSIAETFTHROTAFLERVLNFPQSNL	199	

142	Db	MGTWDPFNPIVITDSDAAFRQQAQVA-MEKARSQI-----HLDSEYKLLLEQITHYQDSEPS	195
200	Qy	CLKREKQDECSGLTQALPSELUKSGADNVSLTGAVSLSASMLTEIFLLQQAQGM-----EPQWG	257
196	Db	C-----KEKHQCSLDAKDTFSANTQQBFGVGQPLKVGNSLVDAFTLQYFEGFPMQDVANG	251
258	Qy	RITDSHOWNTILSLHNAQFYLLQRTPEVARSGRATPLLDLIMAAATLPHHPQQAQYGVTLPT	317
252	Db	GIHTRQWKVLSKLKNGVQDSLFTSPTVARNVAAPLVKVIKVLV-----AERVSAP-	303
318	Qy	SVLPIAGHDNTLANLGCALSLN-WTLPCOPDNTPPGCELVFERWRRLSDNSQIWIOVSLVF	376
304	Db	KVTVLVGHSDNSIASLLTALDFKPYQLHDQYERTPIGGQLVFORHWDGNARDLMIKEYVY	363
377	Qy	QTIQOQMRDKTPLSLNTPPGVEVKLTLACGEERNAQGMCSLAGFTQIUNEA	425
364	Db	QSARQLRNAEALTLKSPAQRVTLKLCPCV-DANGFCFLDKFDNMVMTA	411

[illegible]

A;Title: Human prostatic acid phosphatase: cDNA cloning, gene mapping and protein sequencing
 A;Reference number: A32419; MUID:89228054; PMID:2712834
 A;Accession: A32419
 A;Molecule type: mRNA
 A;Residues: 1-14, 'A', '16', 'ASC', '20', 'CF', '23', 'C', '25-94', 'D', '96-115', 'R', '117-214', 'S', '216-293', 'T'
 A;Cross-references: GB:M24902; NID:g189618; PIDN:AAA60022.1; PID:g189619
 A;Note: the authors translated the codons GAC for residue 95 as Glu, CGT for residue 116 as Arg, P.G.; Govindan, M.V.; Patel, P.C.
 Nucleic Acids Res. 18, 4928, 1990
 A;Title: Nucleotide sequence of human prostatic acid phosphatase determined from a full-length cDNA
 A;Reference number: S11147; MUID:90370491; PMID:2395659
 A;Accession: S11147
 A;Status: preliminary; translation not shown
 A;Molecule type: mRNA
 A;Residues: 1-14, 'AFAC', '20', 'CF', '23', 'C', '25-45', 'H', '47-65', 'RIMPTHPA', '74-138', 'E', '140-156', 'R'
 A;Cross-references: EMBL:X33605; NID:g35683; PIDN:CAA37673.1; PID:g35684
 R;Banas, B.; Blaschke, D.; Fittler, F.; Hoerz, W.
 submitted to the EMBL Data Library, April 1993
 A;Description: Characterization of the promoter of the human prostatic acid phosphatase gene
 A;Reference number: S38863
 A;Accession: S38863
 A;Molecule type: DNA
 A;Residues: 1-40 <BAN>
 A;Cross-references: EMBL:X71391; NID:g416530; PIDN:CAA50514.1; PID:g416531
 R;Virkkunen, P.H.; Hedberg, P.; Palvimo, J.J.; Birr, E.; Porvari, K.; Taavitsainen, P.;
 submitted to the EMBL Data Library, September 1993
 A;Description: Structural organization of human and rat prostate-specific acid phosphatase gene in the human gene promoter.
 A;Reference number: S41251
 A;Accession: S41251
 A;Molecule type: DNA
 A;Residues: 1-40 <VIR>
 A;Cross-references: EMBL:X74961; NID:g439657; PIDN:CAA52913.1; PID:g439658
 R;Lee, H.; Chu, T.M.; Li, S.S.L.; Lee, C.
 Biochem. J. 277, 759-765, 1991
 A;Title: Homodimer and heterodimer subunits of human prostate acid phosphatase.
 A;Reference number: S17042; MUID:91336999; PMID:1908222
 A;Accession: S17042
 A;Status: preliminary
 A;Molecule type: protein
 A;Residues: 33-49 <LEE>
 R;Banas, B.; Blaschke, D.; Fittler, F.; Hoerz, W.
 Biochim. Biophys. Acta 1217, 188-194, 1994
 A;Title: Analysis of the promoter of the human prostatic acid phosphatase gene.
 A;Reference number: S42730; MUID:94153995; PMID:8110833
 A;Accession: S42730
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-40 <BA2>
 A;Cross-references: GB:X71391; NID:g416530; PIDN:CAA50514.1; PID:g416531
 R;Sharief, F.S.; Li, S.S.
 Biochem. Mol. Biol. Int. 33, 561-565, 1994
 A;Title: Nucleotide sequence of human prostatic acid phosphatase ACP gene.
 A;Reference number: I37175; MUID:95038536; PMID:7951074
 A;Accession: I37175
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-386 <RES>
 A;Cross-references: EMBL:U07097; NID:g515995; PIDN:AA860640.1; PID:g515997
 C;Comment: This protein is synthesized under androgen regulation by epithelial cells of prostate
 C;Genetics:
 A;Gene: GDB:ACPP
 A;Cross-references: GDB:I19644; OMIM:171790
 A;Map position: 3q21.3-3q25.2
 A;Introns: 40/3; 72/3; 101/3; 152/3; 185/3; 216/3; 260/3; 288/3; 323/2
 C;Function:
 A;Description: catalyzes the hydrolysis of a wide range of phosphate esters
 C;Superfamily: mammalian acid phosphatase
 C;Keywords: glycoprotein; phosphohistidine; phosphoprotein; phosphoric monoester hydrolase
 F;1-32/Domain: signal sequence #status predicted <SIG>
 F;33-386/Product: acid phosphatase ACP #status experimental <MAT>
 F;43,86/Active site: Arg #status predicted
 F;44/Active site: His (phosphohistidine intermediate) #status predicted

F;94,220,333/Binding site: carbohydrate (Asn) (covalent) #status experimental
 F;161-372,215-313,347-351/Disulfide bonds: #status experimental

Query Match	5.1%	Score 117;	DB 1;	Length 386;
Best Local Similarity	19.8%	Pred. No. 0.14;		
Matches	87;	Conservative 63;	Mismatches 178;	Indels 112; Gaps 15;
QY	3	AILIPFLSLILPLTPOSAPAQSEPELKLRSVIVSRHGVRAPTKATQALQMDVTPDAMPT- 61		
DB	13	SLSLGLFLFLFFWLDRLSLAK-----ELKFTLVFRHGDRSPI-----DTFPTD 56		
QY	62	-----WPVKLGLMTPRGGELIAYLGHYORQRLVADGLLAKKCCPOS---GQVAILADVD 112		
DB	57	PIKESWPQGGTQGLTQGMEOHYELGEYIRKY-----RKFLNESYKHEQYIIRSTDV 109		
QY	113	ERTRTKGEAFAGLAPDCAITVHTQADTSSPDPLFNPLKTGVQCQLDNANVTDAILSRAGG 172		
DB	110	DRTLMSAMTNLAALFPPEGVSIW-----NPILLWOPIPVHTVPLSE----- 150		
QY	173	SIADFTGHRQTAFRELRVINFPPQSNLCKREKQDESCSLTQALPSELKVSADNVSLTGA 232		
DB	151	-----DQLLYLPFRN-CPRFOLESETLKEEFQKRLHPYKDFIATLKG 193		
QY	233	VS-----LASMLTEIFLLOQAQMPGPGWGRITDSHOWNTILSLHNAQFVILLQRTP 283		
DB	194	LSGLHQDLFGIWSKYDPLYCESVHNFTLPSWATEDTWTKLRELSLSLYGIHKQK 253		
QY	284	EVARSRATPLLDLIMAAALTPHPPOKQAYGVTLPTSVLFTIAGHDNLNLAGGAL-NTWL 342		
DB	254	EKSRLOGGVIVNEILNHM-----KRTQIPSYKKLIMYSAHDTTSGLOMALDVYNGLL 307		
QY	343	PGQPDNTPPGELVFERWRRLSDNSQMIQVSLVFQTLQOMRDKTPLSLNTPPGEVKLTLA 402		
DB	308	P--PYASCHLTLYFEKGEYF-----VEMYRN-----ETQHEFPYPLMLP 345		
QY	403	GCEERNAQGMCSLAGFTQIV 422		
DB	346	GCSP-----SCPLERFAELV 360		

Search completed: March 11, 2005, 23:08:35

Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 11, 2005, 22:58:05 ; Search time 172 Seconds
(without alignments)
1309.970 Million cell updates/sec

Title: US-10-034-985-2
Perfect score: 2302
Sequence: 1 MKAILPFLSLIPLTPQSA.....IVNEARIPACLSLRSHHHHH 440

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2235	97.1	432	1	PPA_ECOLI	P07102 escherichia
2	2231	96.9	432	2	Q8GN88	Q8gn88 escherichia
3	2229	96.8	432	2	Q8RKD6	Q8rkd6 escherichia
4	2229	96.8	432	2	Q8RKD7	Q8rkd7 escherichia
5	2229	96.8	432	2	Q8RKD8	Q8rkd8 escherichia
6	2229	96.8	432	2	Q8RKD9	Q8rkd9 escherichia
7	2228	96.8	432	2	Q8RKD5	Q8rkds escherichia
8	2225	96.7	432	2	Q83RW2	Q83rw2 shigella fl
9	2222	96.5	442	2	Q83RW2	Q83rw2 shigella fl
10	2221	96.5	432	2	Q6RK08	Q6rk08 escherichia
11	2218	96.4	432	2	Q7UD08	Q7ud08 shigella fl
12	2213	96.1	446	2	Q8CW75	Q8cw75 escherichia
13	2210	96.0	434	2	Q7AFW3	Q7afw3 escherichia
14	2210	96.0	444	2	Q8XC29	Q8xc29 escherichia
15	1374.5	59.7	433	2	Q676V7	Q676v7 citrobacter
16	1077	46.8	444	2	Q6U677	Q6u677 obesumbacte
17	1054.5	45.8	444	2	Q6TAQ8	Q6taq8 obesumbacte
18	943.5	41.0	441	2	Q66R93	Q66r93 yersinia ps
19	943.5	41.0	441	2	Q8ZFP6	Q8zfp6 yersinia ps
20	545	23.7	413	2	Q8XBZ6	Q8xbz6 escherichia
21	541	23.5	417	1	AGP_PRORE	P19926 providencia
22	539	23.4	413	1	AGP_ECOLI	Q52309 escherichia
23	538	23.4	413	2	Q7UD02	Q7ud02 shigella fl
24	538	23.4	421	2	Q83RV6	Q83rv6 shigella fl
25	529	23.0	413	1	AGP_SALTY	O33921 salmonella
26	526	22.8	414	2	Q9AAQ4	Q9aaq4 caulobacter
27	524	22.8	413	2	Q8Z7P1	Q8z7p1 salmonella
28	523	22.7	413	2	Q6EV19	Q6ev19 enterobacte
29	505	21.9	392	2	Q8P7P6	Q8pp76 xanthomonas
30	505	21.9	443	2	Q8P330	Q8p330 xanthomonas
31	480.5	20.9	435	2	Q8PPF53	Q8ppf53 xanthomonas

32	467.5	20.3	435	2	Q6CZP4	O6czf4 erwinia car
33	433.5	18.8	428	2	Q8GD20	Q8gd20 pseudomonas
34	392	17.0	318	2	Q8CW72	Q8cw72 escherichia
35	373	16.2	414	2	Q8PPX6	Q8ppx6 xanthomonas
36	359	15.6	421	2	Q84CN9	Q84cn9 klebsiella
37	353	15.3	421	2	Q7WSY1	Q7wsy1 klebsiella
38	221.5	9.6	419	2	Q8VQS2	Q8vq2 klebsiella
39	195	8.5	426	2	Q9BZG2	Q9bzg2 homo sapien
40	150	6.5	531	2	Q9GNZ3	Q9gnz3 leishmania
41	140	6.1	423	1	PPAL_HUMAN	P11117 homo sapien
42	140	6.1	542	2	Q9UIA2	Q9uia2 leishmania
43	139.5	6.1	333	2	Q9BZG3	Q9bzg3 homo sapien
44	139	6.0	462	2	Q7QC05	Q7qc05 anopheles g
45	137.5	6.0	406	2	Q66L09	Q66l09 xenopus lae

ALIGNMENTS

RESULT 1
ID_PPA_ECOLI STANDARD; PRT; 432 AA.
AC P07102;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Periplasmic appa protein precursor [Includes: Phosphoanhydride
DE phosphohydrolase (EC 3.1.3.2) (pH 2.5 acid phosphatase) (AP); 4-
DE phytase (EC 3.1.3.26)].
GN Name=appa; OrderedLocusNames=b0980;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
SEQUENCE FROM N.A., AND SEQUENCE OF 23-29.
RC STRAIN=K12;
RX Dassa J., Marck C., Boquet P.L.;
RA "The complete nucleotide sequence of the Escherichia coli gene appa
RT reveals significant homology between the pH 2.5 acid phosphatase and
RT glucose-1-phosphatase".
RL J. Bacteriol. 172:5497-5500 (1990).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503; DOI=10.1126/science.277.5331.1453;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12";
RL Science 277:1453-1474 (1997).
RN [3]
SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,
Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
Yano M., Horiuchi T.;
RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 12.7-28.0 min region on the linkage map";
RL DNA Res. 3:137-155 (1996).
RN [4]
SEQUENCE OF 1-112 FROM N.A.
RC MEDLINE=87271766; PubMed=3038201; DOI=10.1016/0300-9084(87)90045-9;
RA Touati E., Danchin A.;
RT "The structure of the promoter and amino terminal region of the pH 2.5
RT acid phosphatase structural gene (appa) of E. coli: a negative control
RT of transcription mediated by cyclic AMP.";


```

FT STRAND 354 362
FT TURN 363 366

Query Match 97.1%; Score 2235; DB 1; Length 432;
Best Local Similarity 99.5%; Pred. No. 6.3e-165;
Matches 430; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKAIIIPFLSLLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
DB 1 MKAIIIPFLSLLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60

QY 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGOVAILIADVDERTRKTGE 120
DB 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGOVAILIADVDERTRKTGE 120

QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180
DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIADFTGH 180

QY 181 QTAFARELERVLNPPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
DB 181 QTAFARELERVLNPPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240

QY 241 EIFLLOQAQMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300
DB 241 EIFLLOQAQMPGPGWGRITDTSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300

QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALLENWTLPGQPDNTPPGGELVFERW 360
DB 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLANLGALLENWTLPGQPDNTPPGGELVFERW 360

QY 361 RRLSDNSQWIOVSLVFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420
DB 361 RRLSDNSQWIOVSLVFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQMCSLAGFTQ 420

QY 421 INVNEARIPACSL 432
DB 421 INVNEARIPACSL 432

RESULT 3
Q8KRD6 PRELIMINARY; PRT; 432 AA.
AC Q8KRD6;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DE 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Periplasmic phosphoanhydride phosphohydrolase.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90368616; PubMed=2168385;
RA Dassa J., Marck C., Boquet P.-L.;
RT "The complete nucleotide sequence of the Escherichia coli gene appA
RT reveals significant homology between pH 2.5 acid phosphatase and
RT glucose-1-phosphatase.";
RL J. Bacteriol. 172:5497-5500 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93054596; PubMed=1429631;
RA Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,
RA Van Etten R.L.;
RT "Overexpression, site-directed mutagenesis, and mechanism of
RT Escherichia coli acid phosphatase.";
RL J. Biol. Chem. 267:22830-22836 (1992).
DR HSP; P07102; 1DKM.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR000560; HisAc_phsphtse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
KW Hydrolase.
SQ SEQUENCE 432 AA; 46971 MW; AF86C41EA6193AC5 CRC64;

RESULT 2
Q8GN88 PRELIMINARY; PRT; 432 AA.
AC Q8GN88;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DE 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE AppA.
GN Name=appa;
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang Z., He J., Yao B., Zhou Y., Chen Y., Yi Y.;
RA Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Chen Y., Zhu Z., Zhang Z., He J.;
RA Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF537219; AAN28334.1; -.
DR HSP; P07102; 1DKM.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR InterPro; IPR000560; HisAc_phsphtse.
DR Pfam; PF00328; Acid_phosphat_A; 1.
DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
SQ SEQUENCE 432 AA; 47055 MW; 5B355D76E7377737 CRC64;

Query Match 96.9%; Score 2231; DB 2; Length 432;
Best Local Similarity 99.3%; Pred. No. 1.3e-164;
Matches 429; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

Query Match 96.8%; Score 2229; DB 2; Length 432;
 Best Local Similarity 99.3%; Pred. No. 1.8e-164;
 Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
 DB 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60

QY 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPCQSQGVAIADVDERTKTGE 120
 DB 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPCQSQGVAIADVDERTKTGE 120

QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
 DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180

QY 181 ROTAFRELERVLNFPQSNCLCKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 DB 181 ROTAFRELERVLNFPQSNCLCKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLT 240

QY 241 EIFLLOAQOGMPFGWGRITDHSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300
 DB 241 EIFLLOAQOGMPFGWGRITDHSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300

QY 301 LTPHPQKQAYGVTLSVLFIAGHDTNLNGLGALLENWTLPGQPDNTPPGGELVFERW 360
 DB 301 LTPHPQKQAYGVTLSVLFIAGHDTNLNGLGALLENWTLPGQPDNTPPGGELVFERW 360

QY 361 RRLSDNSQWITQVSLVFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
 DB 361 RRLSDNSQWITQVSLVFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

QY 421 INVNEARIPACSL 432
 DB 421 INVNEARIPACSL 432

RESULT 4

Q8RKD7 ID Q8RKD7 PRELIMINARY; PRT; 432 AA.

AC Q8RKD7; 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Periplasmic phosphoanhydride phosphohydrolase.

GN Name=appa;

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90368616; PubMed=2168385;

RA Dassa J., Marck C., Boquet P.-L.L.;

RT "The complete nucleotide sequence of the Escherichia coli gene appA

RT reveals significant homology between pH 2.5 acid phosphatase and

RT glucose-1-phosphatase.";

RL J. Bacteriol. 172:5497-5500(1990).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=9054596; PubMed=1429631;

RA Ostalin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,

RA Van Etten R.L.;

RT "Overexpression, site-directed mutagenesis, and mechanism of

RT Escherichia coli acid phosphatase.";

RL J. Biol. Chem. 267:22830-22836(1992).

DR EMBL; L03373; AAA00005.1; --

DR HSP; P07102; IDKM

DR GO; GO:0003993; F:acid phosphatase activity; IEA.

DR GO; GO:0016787; F:hydrolase activity; IEA.

DR InterPro; IPR00560; HisAc phosphatase.

DR Pfam; PF00328; Acid_phosphat_A; 1.

DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase.
 SQ SEQUENCE. 432 AA; 46971 MW; 755D5E4B1AD916A6 CRC64;

Query Match 96.8%; Score 2229; DB 2; Length 432;
 Best Local Similarity 99.3%; Pred. No. 1.8e-164;
 Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60
 DB 1 MKAILPFLSLLIPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQMLQMDVTPDAMP 60

QY 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPCQSQGVAIADVDERTKTGE 120
 DB 61 TWPVKLGWLTTPRGELIAYLGHYQRLVADGLLAKKGCPCQSQGVAIADVDERTKTGE 120

QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
 DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180

QY 181 ROTAFRELERVLNFPQSNCLCKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 DB 181 ROTAFRELERVLNFPQSNCLCKREKQDECSLTQALPSELKVSADNVSLTGAVSLASMLT 240

QY 241 EIFLLOAQOGMPFGWGRITDHSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300
 DB 241 EIFLLOAQOGMPFGWGRITDHSQWNTLLSLHNAQFYLLQRTPEVARSRAATPLLDLIMAA 300

QY 301 LTPHPQKQAYGVTLSVLFIAGHDTNLNGLGALLENWTLPGQPDNTPPGGELVFERW 360
 DB 301 LTPHPQKQAYGVTLSVLFIAGHDTNLNGLGALLENWTLPGQPDNTPPGGELVFERW 360

QY 361 RRLSDNSQWITQVSLVFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420
 DB 361 RRLSDNSQWITQVSLVFTLQOMRDKTPLSLNTPPGEVKLTLAGCEERNAQGMCSLAGFTQ 420

QY 421 INVNEARIPACSL 432
 DB 421 INVNEARIPACSL 432

RESULT 5

Q8RKD8 ID Q8RKD8 PRELIMINARY; PRT; 432 AA.

AC Q8RKD8; 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Periplasmic phosphoanhydride phosphohydrolase.

GN Name=appa;

OS Escherichia coli.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID=562;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90368616; PubMed=2168385;

RA Dassa J., Marck C., Boquet P.-L.L.;

RT "The complete nucleotide sequence of the Escherichia coli gene appA

RT reveals significant homology between pH 2.5 acid phosphatase and

RT glucose-1-phosphatase.";

RL J. Bacteriol. 172:5497-5500(1990).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=9054596; PubMed=1429631;

RA Ostalin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,

RA Van Etten R.L.;

RT "Overexpression, site-directed mutagenesis, and mechanism of

RT Escherichia coli acid phosphatase.";

RL J. Biol. Chem. 267:22830-22836(1992).

DR EMBL; L03372; AAA00004.1; --

DR HSP; P07102; IDKM.

DR GO; GO:0003993; F:acid phosphatase activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR InterPro; IPR000560; HisAc phosphatase.
 DR Pfam; PF00328; Acid_phosphat_A; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase.
 SQ SEQUENCE 432 AA; 46971 MW; 9885536857FCCFB5 CRC64;

Query Match 96.8%; Score 2229; DB 2; Length 432;
 Best Local Similarity 99.3%; Pred. No. 1.8e-164;
 Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKAILPFLSLILPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
 DB 1 MKAILPFLSLILPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
 QY 61 TWPVKLGLTPRGELIAYLGHYQRLVADGLLAKKGCPOSQGVAVIADVDERTRKTGE 120
 DB 61 TWPVKLGLTPRGELIAYLGHYQRLVADGLLAKKGCPOSQGVAVIADVDERTRKTGE 120
 QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
 DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
 QY 181 ROTAFRELRLVNFPOSNCLKREKODESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
 DB 181 ROTAFRELRLVNFPOSNCLKREKODESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
 QY 241 EIFLLQQAQMPGPGWGRITDTSQWNTLLSLHNAQFVLLQRTPEVARSRATPLDLIMAA 300
 DB 241 EIFLLQQAQMPGPGWGRITDTSQWNTLLSLHNAQFVLLQRTPEVARSRATPLDLIMAA 300
 QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLNLGGALELNTLPGQPDNTPPGGELVFERW 360
 DB 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLNLGGALELNTLPGQPDNTPPGGELVFERW 360
 QY 361 RRLSDNSQWTOVSLVFQTLQMRDKTPLSLNTPPEVKLTLACGEERNAQMCSLAGFTQ 420
 DB 361 RRLSDNSQWTOVSLVFQTLQMRDKTPLSLNTPPEVKLTLACGEERNAQMCSLAGFTQ 420
 QY 421 IVNEARIPACSL 432
 DB 421 IVNEARIPACSL 432

RESULT 6
 Q8RKE0 PRELIMINARY; PRT; 432 AA.
 AC Q8RKE0
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Periplasmic phosphoanhydride phosphohydrolase.
 GN Namesapp;
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90368616; PubMed=2168385;
 RA Dassa J., Marck C., Boquet P.-L.L.;
 RT "The complete nucleotide sequence of the Escherichia coli gene appA
 RT reveals significant homology between pH 2.5 acid phosphatase and
 RT glucose-1-phosphatase."
 RL J. Bacteriol. 172:5497-5500 (1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93054596; PubMed=1429631;
 RA Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,
 RA Van Etten R.L.;
 RT "Overexpression, site-directed mutagenesis, and mechanism of
 RT Escherichia coli acid phosphatase."

RL J. Biol. Chem. 267:22830-22836 (1992).
 DR EMBL; L03370; AAA00002.1; -.
 DR HSP; P07102; IDKM.
 DR GO; GO:0003993; F:acid phosphatase activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR InterPro; IPR000560; HisAc phosphatase.
 DR Pfam; PF00328; Acid_phosphat_A; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase.
 SQ SEQUENCE 432 AA; 46971 MW; 5BB0632D4682EABF CRC64;

Query Match 96.8%; Score 2229; DB 2; Length 432;
 Best Local Similarity 99.3%; Pred. No. 1.8e-164;
 Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKAILPFLSLILPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
 DB 1 MKAILPFLSLILPLTPQSAFAQSEPELKLESVIVSRHGVRAPTKATQLMQDVTPDAMP 60
 QY 61 TWPVKLGLTPRGELIAYLGHYQRLVADGLLAKKGCPOSQGVAVIADVDERTRKTGE 120
 DB 61 TWPVKLGLTPRGELIAYLGHYQRLVADGLLAKKGCPOSQGVAVIADVDERTRKTGE 120
 QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
 DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCOLDNANVTDAILSRAGGSIAFTGH 180
 QY 181 ROTAFRELRLVNFPOSNCLKREKODESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
 DB 181 ROTAFRELRLVNFPOSNCLKREKODESCSLTOALPSELKVSADNVSLTGAVSLASMLT 240
 QY 241 EIFLLQQAQMPGPGWGRITDTSQWNTLLSLHNAQFVLLQRTPEVARSRATPLDLIMAA 300
 DB 241 EIFLLQQAQMPGPGWGRITDTSQWNTLLSLHNAQFVLLQRTPEVARSRATPLDLIMAA 300
 QY 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLNLGGALELNTLPGQPDNTPPGGELVFERW 360
 DB 301 LTPHPQKQAYGVTLPTSVLFIAGHDTNLNLGGALELNTLPGQPDNTPPGGELVFERW 360
 QY 361 RRLSDNSQWTOVSLVFQTLQMRDKTPLSLNTPPEVKLTLACGEERNAQMCSLAGFTQ 420
 DB 361 RRLSDNSQWTOVSLVFQTLQMRDKTPLSLNTPPEVKLTLACGEERNAQMCSLAGFTQ 420
 QY 421 IVNEARIPACSL 432
 DB 421 IVNEARIPACSL 432

RESULT 7
 Q8RKD9 PRELIMINARY; PRT; 432 AA.
 AC Q8RKD9
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Periplasmic phosphoanhydride phosphohydrolase.
 GN Name=appA;
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90368616; PubMed=2168385;
 RA Dassa J., Marck C., Boquet P.-L.L.;
 RT "The complete nucleotide sequence of the Escherichia coli gene appA
 RT reveals significant homology between pH 2.5 acid phosphatase and
 RT glucose-1-phosphatase."
 RL J. Bacteriol. 172:5497-5500 (1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93054596; PubMed=1429631;
 RA Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,

RA Van Etten R.L.;
 RT "Overexpression, site-directed mutagenesis, and mechanism of
 RT Escherichia coli acid phosphatase.";
 RL J. Biol. Chem. 267:22830-22836(1992).
 DR EMBL; L03371; AAA00003.1; -.
 DR HSSP; P07102; IDKM.
 DR GO; GO:0003993; F:acid phosphatase activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR InterPro; IPR000560; Hisac_phspase.
 DR Pfam; PF00328; Acid_phosphat_A; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Hydrolase.
 SQ SEQUENCE 432 AA; 47033 MW; 9F29B9DF9C368175 CRC64;
 Query Match 96.88; Score 2228; DB 2; Length 432;
 Best Local Similarity 99.3%; Pred. No. 2.2e-164;
 Matches 429; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 MKAILIPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
 DB 1 MKAILIPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
 QY 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
 DB 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
 QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
 DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
 QY 181 QTAFRELERVLNFPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 DB 181 QTAFRELERVLNFPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 QY 241 EIFLLQQAQMPGPGWGRITDHSQWNTLLSHNAQFYLLQRTPEVARSRATPLDLIMAA 300
 DB 241 EIFLLQQAQMPGPGWGRITDHSQWNTLLSHNAQFYLLQRTPEVARSRATPLDLIKTA 300
 QY 301 LTHPPQKQAYGVTLPTSVLFIAGHDNTLANLGALSNLWTLPGQDPTPPGGLVPERW 360
 DB 301 LTHPPQKQAYGVTLPTSVLFIAGHDNTLANLGALSNLWTLPGQDPTPPGGLVPERW 360
 QY 361 RRLSDNSQWISQVSLVFTLQMRDKTFLSLNTPPGVKLTLAGCEERNAQCMCSLAGFTQ 420
 DB 361 RRLSDNSQWISQVSLVFTLQMRDKTFLSLNTPPGVKLTLAGCEERNAQCMCSLAGFTQ 420
 QY 421 IVNEARIPACSL 432
 DB 421 IVNEARIPACSL 432

RESULT 8
 Q8RKDS PRELIMINARY; PRT; 432 AA.
 AC Q8RKDS;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Periplasmic phosphoanhydride phosphohydrolase.
 GN Name=appa;
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90368616; PubMed=2168385;
 RA Dassa J., Marck C., Boquet P.-L.L.;
 RT "The complete nucleotide sequence of the Escherichia coli gene appA
 RT reveals significant homology between pH 2.5 acid phosphatase and
 RT glucose-1-phosphatase.";
 RL J. Bacteriol. 172:5497-5500(1990).
 RN [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE=90368616; PubMed=1429631;
 RA Ostanin K., Harms E.H., Stevis P.E., Kuciel R., Zhou M.M.,
 RA Van Etten R.L.;
 RT "Overexpression, site-directed mutagenesis, and mechanism of
 RT Escherichia coli acid phosphatase.";
 RL J. Biol. Chem. 267:22830-22836(1992).
 DR EMBL; L03375; AAA00007.1; -.
 DR HSSP; P07102; IDKM.
 DR GO; GO:0003993; F:acid phosphatase activity; IEA.
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR InterPro; IPR000560; Hisac_phspase.
 DR Pfam; PF00328; Acid_phosphat_A; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 KW Hydrolase.
 SQ SEQUENCE 432 AA; 46990 MW; 951F393EA9A1A47C CRC64;
 Query Match 96.7%; Score 2225; DB 2; Length 432;
 Best Local Similarity 99.3%; Pred. No. 3.8e-164;
 Matches 429; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 MKAILIPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
 DB 1 MKAILIPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTDPDAMP 60
 QY 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
 DB 61 TWPVKLGWLTTPRGELIAYLGHYQORLVADGLLAKKGCPSQGVAIADVDERTKTGE 120
 QY 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
 DB 121 AFAAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFTGH 180
 QY 181 QTAFRELERVLNFPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 DB 181 QTAFRELERVLNFPQSNCLCKREKQDESCSLTQALPSELKVSADNVSLTGAVSLASMLT 240
 QY 241 EIFLLQQAQMPGPGWGRITDHSQWNTLLSHNAQFYLLQRTPEVARSRATPLDLIMAA 300
 DB 241 EIFLLQQAQMPGPGWGRITDHSQWNTLLSHNAQFYLLQRTPEVARSRATPLDLIKTA 300
 QY 301 LTHPPQKQAYGVTLPTSVLFIAGHDNTLANLGALSNLWTLPGQDPTPPGGLVPERW 360
 DB 301 LTHPPQKQAYGVTLPTSVLFIAGHDNTLANLGALSNLWTLPGQDPTPPGGLVPERW 360
 QY 361 RRLSDNSQWISQVSLVFTLQMRDKTFLSLNTPPGVKLTLAGCEERNAQCMCSLAGFTQ 420
 DB 361 RRLSDNSQWISQVSLVFTLQMRDKTFLSLNTPPGVKLTLAGCEERNAQCMCSLAGFTQ 420
 QY 421 IVNEARIPACSL 432
 DB 421 IVNEARIPACSL 432

RESULT 9
 Q83RW2 PRELIMINARY; PRT; 442 AA.
 AC Q83RW2;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Phosphoanhydride phosphorylase; pH 2.5 acid phosphatase;
 DE periplasmic.
 GN Name=appa; OrderedLocusNames=SF0982;
 OS Shigella flexneri.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Shigella.
 OX NCBI_TaxID=623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=301 / Serotype 2a;
 RL MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,

RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RT "Genome sequence of *Shigella flexneri* 2a: insights into pathogenicity
 RT through comparison with genomes of *Escherichia coli* K12 and O157.";
 RL Nucleic Acids Res. 30:4432-4441(2002).
 DR EMBL: AE015127; AAN42610.1; -.
 DR HSSP; P07102; 1DK.
 DR GO; GO:0003993; F:acid phosphatase activity; IEA.
 DR InterPro; IPR000560; HisAc_phsphtse.
 DR Pfam; PF00328; Acid_phosphat_A; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; UNKNOWN 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN 1.
 KW Complete proteome.
 SQ SEQUENCE 442 AA; 48183 MW; 20BCE2C454254680 CRC64;

Query Match 96.5%; Score 2222; DB 2; Length 442;
 Best Local Similarity 98.8%; Pred. No. 6.6e-164;
 Matches 427; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 60
 Db 11 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 70
 Qy 61 TWPVKGLWLTFRGGLIAYLGHYQRLVADGLLAKKGCPCQSGQVATIIADVDERTRKTGE 120
 Db 71 TWPVKGLWLTFRGGLIAYLGHYQRLVADGLLAKKGCPCQSAQVATIIADVDERTRKTGE 130
 Qy 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
 Db 131 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILCRAGGSIADFTGH 190
 Qy 181 RQTAFLERELRVLPPOSNCCLKREKQDESCLTQALPSELKVSADNVSLTGAVSLASMLT 240
 Db 191 RQTFERELRVLPPOSNCCLNREKQDESCLTQALPSELKVSADNVSLTGAVSLASMLT 250
 Qy 241 EIFLLQQAQMPGPGWGRIITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
 Db 251 EIFLLQQAQMPGPGWGRIITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 310
 Qy 301 LTHPPQKQAYGVTLPSTVLFIAGHDNTNLNGLGALSNLWTLPGQPDNTPPGGELVFERW 360
 Db 311 LTHPPQKQAYGVTLPSTVLFIAGHDNTNLNGLGALSNLWTLPGQPDNTPPGGELVFERW 370
 Qy 361 RRLSDNSQWITQVSLVFTLQMRDKTPLSLNTPPGVKLTLAGCEERNAQCMCSLAGFTQ 420
 Db 371 RRLSDNSQWITQVSLVFTLQMRDKTPLSLNTPPGVKLTLAGCEERNAQCMCSLAGFTQ 430
 Qy 421 IVNEARIPACSL 432
 Db 431 IVNEARIPACSL 442

RESULT 10
 Q6RK08
 ID Q6RK08 PRELIMINARY; PRT; 432 AA.
 AC Q6RK08
 DT 05-JUN-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Acid phosphatase/phytase 2.
 GN Name=appa2;
 OS *Escherichia coli*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99194564; PubMed=100925250; DOI=10.1006/bbrc.1999.0361;
 RA Rodriguez E., Han Y., Lei X.G.;
 RT "Cloning, sequencing, and expression of an *Escherichia coli* acid
 RT phosphatase/phytase gene (appa2) isolated from pig colon.";

RL Biochem. Biophys. Res. Commun. 257:117-123(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lei X.G.;
 RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY496073; AAR87658.1; -.
 DR GO; GO:0003993; F:acid phosphatase activity; IEA.
 DR InterPro; IPR000560; HisAc_phsphtse.
 DR Pfam; PF00328; Acid_phosphat_A; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_2; UNKNOWN 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_1; 1.
 FT CHAIN 20 432 acid phosphatase/phytase 2.
 SQ SEQUENCE 432 AA; 47041 MW; 71B0E5ED2EA2674 CRC64;

Query Match 96.5%; Score 2221; DB 2; Length 432;
 Best Local Similarity 99.1%; Pred. No. 7.7e-164;
 Matches 428; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 60
 Db 1 MKAILPFLSLLPLTPQSAFAQSEPELKLESVVIVSRHGVRAPTKATQLMQDVTPDAMP 60
 Qy 61 TWPVKGLWLTFRGGLIAYLGHYQRLVADGLLAKKGCPCQSGQVATIIADVDERTRKTGE 120
 Db 61 TWPVKGLWLTFRGGLIAYLGHYQRLVADGLLAKKGCPCQSAQVATIIADVDERTRKTGE 120
 Qy 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
 Db 121 APAAGLAPDCAITVHTQADTSSPDPLFNPLKTGVCQLDNANVTDAILSRAGGSIADFTGH 180
 Qy 181 RQTAFLERELRVLPPOSNCCLKREKQDESCLTQALPSELKVSADNVSLTGAVSLASMLT 240
 Db 181 RQTAFLERELRVLPPOSNCCLNREKQDESCLTQALPSELKVSADNVSLTGAVSLASMLT 240
 Qy 241 EIFLLQQAQMPGPGWGRIITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
 Db 241 EIFLLQQAQMPGPGWGRIITDSHOWNTLLSHNAQFYLLQRTPEVARSRATPLLDLIMAA 300
 Qy 301 LTHPPQKQAYGVTLPSTVLFIAGHDNTNLNGLGALSNLWTLPGQPDNTPPGGELVFERW 360
 Db 301 LTHPPQKQAYGVTLPSTVLFIAGHDNTNLNGLGALSNLWTLPGQPDNTPPGGELVFERW 360
 Qy 361 RRLSDNSQWITQVSLVFTLQMRDKTPLSLNTPPGVKLTLAGCEERNAQCMCSLAGFTQ 420
 Db 361 RRLSDNSQWITQVSLVFTLQMRDKTPLSLNTPPGVKLTLAGCEERNAQCMCSLAGFTQ 420
 Qy 421 IVNEARIPACSL 432
 Db 421 IVNEARIPACSL 432

RESULT 11
 Q7UD08
 ID Q7UD08 PRELIMINARY; PRT; 432 AA.
 AC Q7UD08
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Phosphoanhydride phosphorylase; pH 2.5 acid phosphatase.
 GN Name=appa; OrderedLocusNames=S1048;
 OS *Shigella flexneri*.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Shigella.
 OX NCBI_TaxID=623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=24577;
 RX STRAIN=22590274; PubMed=12704152;
 RX DOI=10.1128/IAI.71.5.2775-2786.2003;
 RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
 RA Fournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
 RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
 RA Schwartz D.C., Blattner F.R.;

DR InterPro; IPR000560; Hisac_phosphatase.
 DR Pfam; PF00328; Acid_phosphatase; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 SQ SEQUENCE 434 AA; 47337 MW; F197DF7D1869F9C4 CRC64;

Query Match 96.0%; Score 2210; DB 2; Length 434;
 Best Local Similarity 98.4%; Pred. No. 5.5e-163;
 Matches 427; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

QY 1 MKAILIPFLLSLPLTPQSAFAQS--EPKLKESVIVSRHGVRAPTKATQMLQMDVTPDA 58
 DB 1 MKAILIPFLLSLPLTPQSAFAQSEPEPELKESVIVSRHGVRAPTKATQMLQMDVTPDA 60
 QY 59 WPTWPKLGLWLTFRGGELIAYLGHYQRLVADGLLAKKGCPOSQGVIAIADYDERTKT 118
 DB 61 WPNWPKLGLWLTFRGGELIAYLGHYQRLVADGLLTKGCPQPGQVIAIADYDERTKT 120
 QY 119 GEAFAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFT 178
 DB 121 GEAFAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFT 180
 QY 179 GHRTAFRELERVLPQSNCLKREKQDESCSLTQALPSELKVSADNVSLSLTGAVSLASM 238
 DB 181 GHRTAFRELERVLPQSNCLKREKQDESCSLTQALPSELKVSADNVSLSLTGAVSLASM 240
 QY 239 LTFIFLLQQAQMGPEPGWGRITDTHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIM 298
 DB 241 LTFIFLLQQAQMGPEPGWGRITDTHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIM 300
 QY 299 AALTTPHPKQKQAVGVTLPSTVLFIAGHTNLANLGGALELNWTLPGQPDNTPPGGELVFE 358
 DB 301 IALTTPHPKQKQAVGVTLPSTVLFIAGHTNLANLGGALELNWTLPGQPDNTPPGGELVFE 360
 QY 359 RWRLSDNSQWISVLFQTLQOMRDKTPLSLNTPPEVKLTLAGCEERNAQMCSLAGF 418
 DB 361 RWRLSDNSQWISVLFQTLQOMRDKTPLSLNTPPEVKLTLAGCEERNAQMCSLAGF 420
 QY 419 TQIVNEARIPACSL 432
 DB 421 TQIVNEARIPACSL 434

RESULT 14
 Q8XC29 PRELIMINARY; PRT; 444 AA.
 AC Q8XC29;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
 DE Phosphoanhydride phosphorylase; pH 2.5 acid phosphatase;
 DE periplasmic.
 GN Name=appa; OrderedLocusNames=z1397;
 OS Escherichia coli O157:H7.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=83334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927 / EHEC;
 RX MEDLINE=21074935; PubMed=11206551; DOI=10.1038/35054089;
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";
 RL Nature 409:529-533 (2001).
 DR ENBL; AE005292; AAG55528.1; --
 DR PIR; D85633; D85633.
 DR PIR; H90770; H90770.
 DR HSSP; P07102; 1DKL.

DR CO; CO:0003993; P:acid phosphatase activity; IEA.
 DR Pfam; PF00328; Acid_phosphatase; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 KW Complete proteome.
 SQ SEQUENCE 444 AA; 48388 MW; A9AA7E6653AF247E CRC64;

Query Match 96.0%; Score 2210; DB 2; Length 444;
 Best Local Similarity 98.4%; Pred. No. 5.7e-163;
 Matches 427; Conservative 0; Mismatches 15; Indels 2; Gaps 1;

QY 1 MKAILIPFLLSLPLTPQSAFAQS--EPKLKESVIVSRHGVRAPTKATQMLQMDVTPDA 58
 DB 11 MKAILIPFLLSLPLTPQSAFAQSEPEPELKESVIVSRHGVRAPTKATQMLQMDVTPDA 70
 QY 59 WPTWPKLGLWLTFRGGELIAYLGHYQRLVADGLLAKKGCPOSQGVIAIADYDERTKT 118
 DB 71 WPNWPKLGLWLTFRGGELIAYLGHYQRLVADGLLTKGCPQPGQVIAIADYDERTKT 130
 QY 119 GEAFAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFT 178
 DB 131 GEAFAGLAPDCAITVHTQADTSSPDPLFNPLKTVGCOLDNANVTDAILSRAGGSIAFT 190
 QY 179 GHRTAFRELERVLPQSNCLKREKQDESCSLTQALPSELKVSADNVSLSLTGAVSLASM 238
 DB 191 GHRTAFRELERVLPQSNCLKREKQDESCSLTQALPSELKVSADNVSLSLTGAVSLASM 250
 QY 239 LTFIFLLQQAQMGPEPGWGRITDTHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIM 298
 DB 251 LTFIFLLQQAQMGPEPGWGRITDTHQWNTLLSLHNAQFYLLQRTPEVARSRAATPLDLIM 310
 QY 299 AALTTPHPKQKQAVGVTLPSTVLFIAGHTNLANLGGALELNWTLPGQPDNTPPGGELVFE 358
 DB 311 IALTTPHPKQKQAVGVTLPSTVLFIAGHTNLANLGGALELNWTLPGQPDNTPPGGELVFE 370
 QY 359 RWRLSDNSQWISVLFQTLQOMRDKTPLSLNTPPEVKLTLAGCEERNAQMCSLAGF 418
 DB 371 RWRLSDNSQWISVLFQTLQOMRDKTPLSLNTPPEVKLTLAGCEERNAQMCSLAGF 430
 QY 419 TQIVNEARIPACSL 432
 DB 431 TQIVNEARIPACSL 444

RESULT 15
 Q676V7 PRELIMINARY; PRT; 433 AA.
 AC Q676V7;
 DT 25-OCT-2004 (TReMBLrel. 28, Created)
 DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
 DE Phytase.
 GN Name=phyA;
 OS Citrobacter freundii.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Citrobacter.
 OX NCBI_TaxID=546;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VKPM B-4090;
 RA Zinin N.V., Syneokii S.P.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY390262; AAR89622.1; --
 DR InterPro; IPR000560; HisAc_phosphatase.
 DR Pfam; PF00328; Acid_phosphatase; 1.
 DR PROSITE; PS00616; HIS_ACID_PHOSPHAT_1; 1.
 DR PROSITE; PS00778; HIS_ACID_PHOSPHAT_2; UNKNOWN_1.
 SQ SEQUENCE 433 AA; 48506 MW; A7923288F2FCFD44 CRC64;

Query Match 59.7%; Score 1374.5; DB 2; Length 433;
 Best Local Similarity 61.3%; Pred. No. 3.8e-98;
 Matches 265; Conservative 51; Mismatches 115; Indels 1; Gaps 1;

